DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93943

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

MARINE OFFICER ATTRITION MODEL

by

William John Esmann

June 1984

Thesis Advisor: Dan C. Boger

Approved for Public Release; Distribution Unlimited

T217399

DUDLEY KNOX LIBR
NAVAL PO
MONTEREY, CA 9343

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
Marine Officer Attrition Model

5. TYPE OF REPORT & PERIOD COVERED
Master's Thesis; June, 1984

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
William John Esmann

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93943

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93943

12. REPORT DATE
June, 1984

13. NUMBER OF PAGES
72

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; Distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Marine Officer Attrition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Predicting officer attrition is a major difficulty in the Planning, Programming and Budgeting Process. The Marine Corps until recently used an averaging method of determining out-year attrition. The purpose of this paper is to apply econometric techniques to the problem of predicting attrition. This paper develops a simple model which enables the unskilled user to accurately predict attrition.

Approved for public release; distribution unlimited.

Marine Officer Attrition Model

by

William John Esmann
Major, United States Marine Corps
B.S., U.S. Military Academy, 1970

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 1984

ABSTRACT

Predicting officer attrition is a major difficulty in the Planning, Programming, and Budgeting Process. The Marine Corps until recently used an averaging method of determining out-year attrition. The purpose of this paper is to apply econometric techniques to the problem of predicting attrition. This paper develops a simple model which enables the unskilled user to accurately predict attrition.

TABLE OF CONTENTS

I. INTRODUCTION 8
A. GENERAL 8
B. BACKGROUND 8
1. Determination of Variables 8
2. Uses of the Model 10
3. Choice of Regression 10

II. STRUCTURE OF THE DATA BASE 12
A. ATTRITION DATA 12
B. ECONOMIC DATA 13

III. INITIATIVE DEVELOPMENT OF THE MODEL 17
A. INTERVIEWS 17

IV. METHODOLOGY AND ANALYSIS 22
A. DATA ANALYSIS 22
1. Scatter plot analysis 22
2. Description of Variables 25
3. Promotion Potential 26
B. BASIC REGRESSION MODEL 27
C. ELEMENTS OF THE SAS OUTPUT 27
1. F value 29
2. R-square 31
3. t-test 31
4. Coefficient of Variation 32
5. Prediction Interval 32
6. Autocorrelation 32
7. Durbin-Watson Test 32
D. EFFECTS OF SAMPLE SIZE ON REGRESSION 33
E. DIFFICULTIES WITH AUTOCORREIATION 34

F. VERIFICATION OF REGRESSION ASSUMPTIONS . . . 36
G. MODEL VARIABLES AND COEFFICIENTS 36
H. APPLICATION OF STATISTICAL TESTS TO THE MODEL 36
1. Lieutenant Colonels 38
2. Major 40
3. Captain 42
4. Lieutenants 43

V. RESULTS AND CONCLUSIONS 44
A. MODEL COMPARISONS 44
B. PREDICTION INTERVALS 45
C. SENSITIVITY ANALYSIS 46
D. CONCLUSION 47

APPENDIX A: SCATTER PLOTS AND COMPARISON DIAGRAMS . . . 49

APPENDIX B: GENERAL DIFFERENCING EQUATIONS 60
A. MODEL EQUATIONS RESULTING FROM GENERAL DIFFERENCING PROCEDURES. 60
1. Lieutenant ground. 60
2. Captain total. 60
3. Major aviation. 60
4. Major ground. 60
B. COMPARISON OF ORIGINAL 61

APPENDIX C: PREDICTION INTERVALS 62

LIST OF REFERENCES 70

BIBLIOGRAPHY 72

INITIAL DISTRIBUTION LIST 73

LIST OF TABLES

I. MARINE OFFICER ATTRITION 14
II. ECONOMIC FACTORS 15
III. RESULTS OF GENERAL DIFFERENCING PROCEDURE 37
IV. MODEL VARIABLES AND COEFFICIENTS 38
V. STATISTICAL ANALYSIS OF THE MODEL 39
VI. COMPARISON OF MODEL RESULTS 45
VII. SENSITIVITY ANALYSIS OF 1984 MODEL 47
VIII. RESULTS OF GENERAL DIFFERENCING 61

LIST OF FIGURES

4.1	Lieutenant Colonel Ground	24
4.2	Example of SAS Output	28
4.3	Example of a SAS Program	30
5.1	Prediction Intervals	46
A.1	Total Lieutenant Colonels	50
A.2	Total Majors	51
A.3	Total Captains	52
A.4	Ground Lieutenant Colonels	53
A.5	Ground Majors	54
A.6	Ground Captains	55
A.7	Ground Lieutenants	56
A.8	Aviation Lieutenant Colonels	57
A.9	Aviation Majors	58
A.10	Aviation Captains	59
C.1	Lieutenant Colonel Ground	62
C.2	Lieutenant Colonel Aviation	63
C.3	Lieutenant Colonel Total	63
C.4	Major Ground	64
C.5	Major Ground/Using General Differencing	64
C.6	Major Aviation	65
C.7	Major Air/Using General Differencing	65
C.8	Major Total	66
C.9	Captain Ground	66
C.10	Captain Aviation	67
C.11	Captain Total	67
C.12	Captain Total/Using General Differencing	68
C.13	Lieutenant Ground	68
C.14	Lieutenant Ground/Using General Differencing	69

I. INTRODUCTION

A. GENERAL

This paper results from a Headquarters Marine Corps (HQMC) request for an officer attrition model based on economic factors. The model will replace the former model of an average of previous years' attrition. The model will be used for budget and promotion planning and will satisfy the need for more accuracy in these projections.

Previous studies in this area have suggested that economic factors are chiefly responsible for officer attrition. The purpose of this paper was to expand on a Center for Naval Analysis (CNA) study [Ref. 1] and develop a regression model which is specific to each officer grade and the components within that grade, such as aviation and ground. The model will produce an attrition rate which, when applied to an average annual officer strength based on total manhours, would give a prediction of attrition for the following fiscal year.

B. BACKGROUND

1. Determination of Variables

In determining which independent variables to use in developing this model, a great deal of thought went into deciding exactly what makes an officer leave the Marine Corps. A variety of ideas were discussed with officers ranging in rank from lieutenants to generals. Studies were also made of informal surveys of former officers who returned questionnaires relating to their decision to resign. The results of this research narrowed choices of

variables to three categories: military pay, the economy, and promotion potential. Promotion potential indices are being developed at this time which will give an indication of an officer's potential for advancement. These indices could be used in a binary choice attrition model in the future. Except at very low grade levels pay, a point of dissatisfaction with many officers, did not correlate strongly with an officer's decision to leave the service when used as the independent variable in a linear regression on attrition. In order to test the influence of other economic variables on attrition, a variety of variables were developed that would indicate trends in economic activity. Among them were managerial unemployment, professional technical unemployment, consumer price index, and GNP. Data for these variables were obtained from the Department of Commerce and library research. A ratio of civilian to military pay developed by the Center for Naval Analyses was also used [Ref. 2]. Most of these variables are self-explanatory, but professional-technical unemployment is identified by the Department of Commerce as unemployment among lawyers, pilots, computer specialists, teachers, programmers, etc. [Ref. 3].

The primary difficulty encountered in doing this project was obtaining sufficient data points with which to run a meaningful regression. The reason for this difficulty was the state of the Marine Corps automated data reporting system prior to 1976. During the period 1970 to 1976 the Marine Corps was instituting its first automated personnel reporting system. Many difficulties were encountered during this period and, as a result, data from this period is extremely unreliable. Data at the Defense Manpower Data Center is based on input from HQMC and after careful study was found to have similar problems. Add to this the attrition problems of the Vietnam War, and the problems in

extending the data base were insurmountable for the purposes of this paper. The Marine Corps system was amended in 1976 and, by 1977, the system was reporting data with over 95 percent accuracy.

The statistical software programming system used in building the model was the SAS Institute's Statistical Analysis System [Ref. 4]. Problems with autocorrelation were identified and resolved using the Hildreth-Lu procedure. [Ref. 5]. Model fits were exceptionally good by F statistic and R-squared standards. The model is currently being used by HQMC to determine officer losses for FY 84 and will serve as the basis for further development into an expanded model.

2. Uses of the Model

The predictions from this model will be used in several ways. The first, which is being prepared at this moment, is the prediction of expected manpower levels for FY 1984. This is an annual process in which expected attrition determines expected accession requirements and, hence, officer recruiting goals. From these figures the manpower budget is then determined.

The preparation of promotion zones is another use of this model. With an accurate prediction of attrition, eligibility zones can be determined well in advance thus facilitating Marine Corps planning as well as officers' personal planning. Other areas affected by this model will be school quotas and retention bonuses, such as Aviation Officer's Continuation Pay.

3. Choice of Regression

The primary reasons for choosing a regression model in this case were simplicity and the sparsness of data. The model will be used by mathematically unsophisticated

officers who have little or no desire to manipulate complex mathematical expressions or to do extensive computer work. With a regression model, answers are provided which express an easily understood relationship between a cause and an effect, which in this case are unemployment and officer attrition rates.

Factors for an ACOI-type model were not developed because of an intuitive error in this type model as it applies to Marine Corps officers. The ACOL type models relate voluntary attrition to variations in military and civilian compensation. Essentially, the ACOL-type models say an individual will leave the military if he senses an erosion in his present compensation in relation to civilian compensation which presages an erosion of future benefits [Ref. 6]. However, Marine Officer motivation for continued service is not based on monetary rewards as much as it is a variety of other factors such as patriotism, pride in service, and a basic satisfaction with his standard of living that is acceptable given an opportunity to continue in service. The ACOL model presumes that an officer is a reasonable man in the legal or economic sense, and that he will weigh the financial benefits of military service versus civilian life, and whichever becomes more favorable will be his career of choice. In this author's experience Marine Officers make an emotional commitment to service and tend to remain in service until they have reached their goals or have determined that they no longer have a possibility of reaching them. This intuition was confirmed by regressions of attrition on civilian to military pay ratios which show little explanatory power.

II. STRUCTURE OF THE DATA BASE

A. ATTRITION DATA

It was necessary to first explore alternative sources of data before beginning the analysis because of the problems mentioned in Chapter One concerning the Marine Corps official data base. The Defense Manpower Data Center (DMDC) maintains data files on all DOD personnel, both active and inactive. From these files DMDC is capable of extracting data on the number of officers and their grade attriting in any one year. Files on Marine Corps officers extend back to 1971 and include unrestricted officers as well as limited duty officers. Data submitted to DMDC by the Marine Corps were obtained from the Marine Corps personnel reporting system. The difficulties with the data were a result of the reporting system prior to 1977. Attrition data were summerized on a semi-annual basis by HQMC and sent via tape to DMDC. Since the reporting system at this time was approximately six months in arrears, the problems of determining exactly when a Marine left the service were difficult to resolve. Because of the problems in the HQMC data it was expected that similar problems would occur in the DMDC data, but an attempt was made to resolve these difficulties. By taking an overall list of attrition from 1971 to 1983 it was hoped that the data could be reduced to annual attrition data by summarizing the data by separation dates. The results obtained from this effort were distinctly different from HQMC data for the same period. Part of the difference can be attributed to the Department of Defense (DOD) coding system which identifies the reasons for an individual's attrition and part to the residual effects of the Vietnam

War. Further compounding the problem were instances of lost tapes and duplication of names. Differences between DOD and Marine Corps coding systems caused additional problems of differentiation between unrestricted officers and Limited Duty Officers. The result of these data problems was to disallow the use of attrition data from years prior to 1977.

The data actually used were obtained from HQMC (Officer Plans Section) and dates from Fiscal Year 1977 through Fiscal Year 1983. To ensure accuracy of this data, an intensive effort was made by the Officer Plans Section to verify the data by comparison with data maintained independently of the personnel system. The data are in the form of total attrition and average strength for a given fiscal year. The average strength is computed by using man-hours totaled over the year and divided by the number of days in a year. From this information the annual attrition rates were then computed. These data are listed in Table I by rank, component, and year.

E. ECONOMIC DATA

Economic data were obtained primarily from the Statistical Office of the Department of Commerce. In particular, all unemployment information was provided by this source. The data were available in several forms including the raw number of unemployed and the total number in the work category, as well as percentages of unemployed in each work category.

Unemployment data appear in a large number of categories reflecting the enormous variety of occupations in this country. Previous studies on this subject by CNA used only the figures for gross unemployment of males over age 18 [Ref. 7]. Intuitively it was not reasonable that this measure would accurately reflect attrition of Marine

TABLE I

MARINE OFFICER ATTRITION

RANK	1977	1978	1979	1980	1981	1982	1983
Ground							
LTCOL	.115	.148	.137	.116	.111	.096	.094
MAJOR	.086	.089	.093	.084	.087	.044	.062
CAPT	.073	.088	.092	.104	.092	.090	.076
1STLT	.225	.233	.210	.174	.157	.147	.143
Aviation							
LTCOL	.139	.172	.144	.121	.125	.122	.060
MAJOR	.071	.079	.059	.052	.048	.034	.036
CAPT	.129	.176	.179	.156	.148	.091	.100
1STLT	.067	.069	.069	.034	.032	.025	.019
Total							
LTCOL	.122	.155	.138	.119	.117	.105	.086
MAJOR	.078	.085	.080	.072	.072	.040	.047
CAPT	.096	.123	.124	.122	.112	.090	.093
1STLT	.166	.176	.147	.123	.109	.102	.099

Officers who are primarily leaders, or in civilian terms, managers. Those officers who do not fall in the managerial category are, for the most part, technically oriented or pilots. For this reason the author examined certain subcategories of unemployment such as managerial/administrative, aviation, and professional/technical. These categories are defined in detail in Chapter Four below and relate closely to the types of work for which Marine Officers are qualified. The economic data showing significant relation to Marine Officer attrition based on linear

TABLE II

ECONOMIC FACTORS

Year	UN	UN1	PT	PT1	PAY	GNP	CPI
1977	.028	.031	.030	.032	1.036	5.3	6.5
1978	.021	.028	.026	.030	1.024	4.4	7.6
1979	.021	.021	.024	.026	1.000	2.3	11.5
1980	.024	.021	.025	.024	1.000	-.2	13.5
1981	.027	.024	.028	.025	1.050	-.3	10.2
1982	.036	.027	.033	.028	1.045	-.8	6.0
1983	.035	.036	.031	.033	1.022	N/A	N/A

regression analysis are listed in Table II under the variable names UN (managerial administrative) and PT (professional technical). The variables UN1 and PT1 in Table II are the variables UN and PT lagged by one year.

Research at the Pentagon library and at the Naval Postgraduate School provided additional information on unemployment categories, as well as data on civilian and military pay. The actual data on pay used in the model was obtained from a Center for Naval Analyses study [Ref. 8]. The study, by Kathleen Utgoff, computed pay as a ratio of civilian and military pay indices. The indices were computed by choosing 1980 as a base year and dividing each year in the sample by the value for 1980. The civilian pay index was computed in the same way using 1980 as the base year. The ratio of these two indices then showed a continuing relationship between civilian and military pay. This data is listed under the variable name PAY in Table II . Data on the gross national product (GNP) and the consumer

price index (CPI) were obtained from library research as well. These variables are also defined in Chapter Four.

III. INTUITIVE DEVELOPMENT OF THE MODEL

A. INTERVIEWS

Procedures followed in developing this model were standard data analysis methods learned at the Naval Postgraduate School. The first step in the process was an intuitive study of the problem in which the analyst determines from background information what factors might have a logical effect on the result he is trying to predict. Since the purpose of this study was to determine what economic factors might influence officer attrition, the search for causes began in this area. To accomplish this step, a series of interviews was held with a variety of senior officers at HQMC, including the officer-in-charge of the Marine Corps Officer Plans Section, the Director of Officer Career planning, and a number of assignment monitors. All of these individuals deal with officer career patterns on a daily basis and all had strong opinions as to the reasons for officer attrition. (It is important to note that these officers were all near retirement age.) Additionally, they provided a variety of other insights as to the officer mind set regarding his career choices. Results from these interviews showed a general disagreement on which particular economic factor was most important in influencing an officer's decision to leave the service. None mentioned unemployment, but most selected pay, benefits, and the economy in general as significant factors in attrition. Aside from these economic factors, one item, the potential for continued promotion, was mentioned by almost everyone. In sum, their feeling was that an officer's success was the determining factor as to whether he left the Marine Corps.

These officers did feel, however, that the economy had some effect on the timing of an officer's departure from the Marine Corps.

The author received several years' worth of questionnaires from the Career Planning section which had been collected from Lieutenant Colonels who had recently resigned. There were approximately 180 of these packages. Reading through the questionnaire answers and the accompanying written narratives provided an invaluable glimpse into the thoughts of men who, at the time of writing, had just made the decision to leave the service. In general, there were two reasons why these men left the service when they did. The first was that they felt the time was right for them to make the choice of continued service until they were either selected, or passed over for Colonel, or resign and pursue a second career. (All were 20-23 year retirees.) Their concern was that staying on in the Marine Corps past age 45 would reduce their chances for obtaining satisfactory employment when they eventually retired. This was based on the feeling that officers from the 42-45 year old age group would have a significantly reduced chance of being hired by a company in which they could continue to progress. The second major factor in the officers' decisions was their opinion that they had reached their promotion limit. With the assumption of minimal potential for promotion, their best financial option was to apply the logic in the previous paragraph and leave the Marine Corps.

Based on the aforementioned interviews and personal experiences, the author began an analysis of the officer as he moves through his career. It was readily apparent that there were key periods in which an officer was most likely to leave the service. These periods were based on promotion points in the service. Under normal circumstances, lieutenants are promoted to Captains in the fifth year of

commissioned service, Captains to Majors during the tenth year, Majors to Lieutenant Colonels during the sixteenth year of service, and Lieutenant Colonels to Colonels during the twenty-second year of service. Integrating this with the fact that retirement benefits are not achieved until the twentieth year of total service yields a brief picture of the promotions in a Marine Officer's career. The following paragraphs give a more detailed discussion of each rank.

Lieutenants normally have a four or five year initial term of service depending on their source of commission and component (air or ground). Essentially, no aviation lieutenants leave the service. This is because of their longer initial term of service, which causes them to reach the grade of Captain before they attrite. Aviation lieutenant attrition is purely a function of accidents, illness, and disciplinary problems and, consequently, was not modeled. Ground lieutenants however, do attrite in significant numbers as the length of service requirements for them are much less severe. Attrition in this grade is based on an individual's analysis of his future. Presumably, the fundamental question is one which results in a choice of a military or civilian life based on his goals and ambitions. Because the perceptions of the attainability of these goals must reasonably lessen during an economic downturn the economy should have a significant effect on an officer's decision to stay or leave. If the economy is depressed it is a difficult time for an officer to leave the security of the military, especially when there is a high probability of extending his service by one of several short term agreements. These agreements can extend a selected officer's service to a more favorable period of time for his exit. Additionally, there is a smaller chance of obtaining a regular commission which would extend his service until he resigns or is passed over for promotion two consecutive times.

For Captains there is a much greater period of flexibility given that they are regular officers, or reserves with an extension of service agreement. Since Captains are promoted at five years of commissioned service and are either passed over for or promoted to Major by their eleventh year of service, an individual may leave the service voluntarily at any point in a period of over six years. Within this period of six years a Captain will have great flexibility in making two key decisions. The first is whether or not to remain in the Marine Corps. The second is when he will leave the Marine Corps if he decides to resign. Once the first is decided, the timing of the second will depend on his ability to obtain satisfactory employment. This of course will depend on the economy as well as his skills.

With the rank of Major the problem is simpler. Pre-DOPMA (Defense Officer Personnel Management Act) Majors have a guaranteed length of service of twenty years commissioned service. Since they are promoted during the tenth and eleventh years of commissioned service and are either passed over or selected for Lieutenant Colonel in their sixteenth to twentieth years of service, they have a period of six to ten years in which to make a decision about leaving. Post-DOPMA Majors do not have a guaranteed length of service of twenty years but there are provisions which allow the retention of these officers on active duty based on their skills and the findings of a special board. Since DOPMA has only affected two year groups it is difficult to see how it will affect the attrition of Majors. By the time officers make the rank of Major most have committed themselves for a twenty year term of service. By reaching the twenty-year retirement point they assure themselves of a very satisfactory retirement program.

At the sixteenth year of service, a Major is first eligible for selection to Lieutenant Colonel. The selection process may be repeated for four years until the officer is either promoted or passed over and retired at twenty years of service. During this later period it is very unlikely that any Major will leave the service unless he has prior enlisted service which would help him achieve twenty years of service early. Lieutenant Colonels probably have the most flexibility of all officers. They are usually promoted at sixteen years of commissioned service and may continue in service until the twenty-sixth year of service. At any point after their third year in grade they may retire although it is considered economically foolish to leave before retirement. From nineteen years of service on, however, the Lieutenant Colonel can pick his time to leave. Most men at this stage of life would not leave unless they had a job either already arranged or the economy was in such a condition that obtaining a job was not difficult.

The result of the above discussion is that Marine Officers make their decisions to leave the service for reasons generally unrelated to the economy. The timing of their decision, however, is directly related to the economy since they invariably require employment soon after their departure from the service. Thus a poor economy will reduce the attrition of Marine officers, while a robust economy will cause attrition to increase given that there is a substantial pool of officers who have made the decision to leave at any given time.

IV. METHODOLOGY AND ANALYSIS

A. DATA ANALYSIS

Data analysis for this project was guided by the intuitive analysis of the problem given in the previous chapter. The analysis pointed this research toward a particular set of variables which were then examined by the author with the intention of determining relationships between them and officer attrition.

1. Scatter plot analysis

The next step in this process was to verify the intuitive relationships between the prospective independent and dependent variables by visual inspection. This was done using scatter plots of the data. A scatter plot is a simple plotting of two-dimensional data on (x,y) coordinates using a specified scale. In this case, attrition data was plotted on the y coordinate versus a variety of economic data on the x coordinate. The resulting set of points should show some type of a pattern if there is a relationship between the two variables. The patterns could have a variety of shapes but in the case of this data they most likely will be linear in form. If they are linear then the hypothesized cause and effect relationship is easy to observe between the two variables. The actual plots were constructed using IBM's experimental graphics software, Grafst3, in the NPS graphics room. Specific functions used were the Scatter Plot Analysis for the scatter plots and General Plot for the linear comparisons.

Scatter plots were run on officer attrition versus the economic variables defined in Chapter II as well as the

rates of change of these versus the rates of change of officer attrition. The most significant relationships occurred between the various forms of unemployment and officer attrition. Also notable was variation in these relationships with respect to officer grade and component. In conjunction with the scatter plots the economic factors and the attrition rates were also plotted on one scale versus each year since 1977. This allowed a direct comparison between the two factors and served to further illustrate the relationship. Figure 4.1 contains a sample of the graphs for Lieutenant Colonel, Ground. Note that a least squares regression line has been included in the scatter to emphasize possible relationships. Only the variables finally chosen for the model were displayed on the plots. The remainder of the plots for the chosen variables are displayed in Appendix A. After comparing the regression fits for all variables, I decided to use the variables UN and UN1 in my regression equations. Additionally the variable PI was lost due to its redefinition by the Department of Commerce in 1983.

Most of these plots show a strong correlation between unemployment in general and Marine Officer attrition. The correlation between aviation officer attrition and unemployment, however, was much weaker. This was caused by the initiation of Aviation Officer Continuation Pay in 1981 which gives a large bonus of as much as $6,000 per year for six years to aviation officers electing to continue in service. The program applied to all ranks provided the individual met certain active duty in flight status requirements. This action by DOD has had its desired effect, and almost 500 officers, are continuing in service (including fifty Lieutenant Colonels who otherwise would not). This accounts, in part, for the dramatic reduction in attrition rates for aviation Lieutenant Colonels in 1983 and for aviation Majors and Captains in 1982 and 1983.

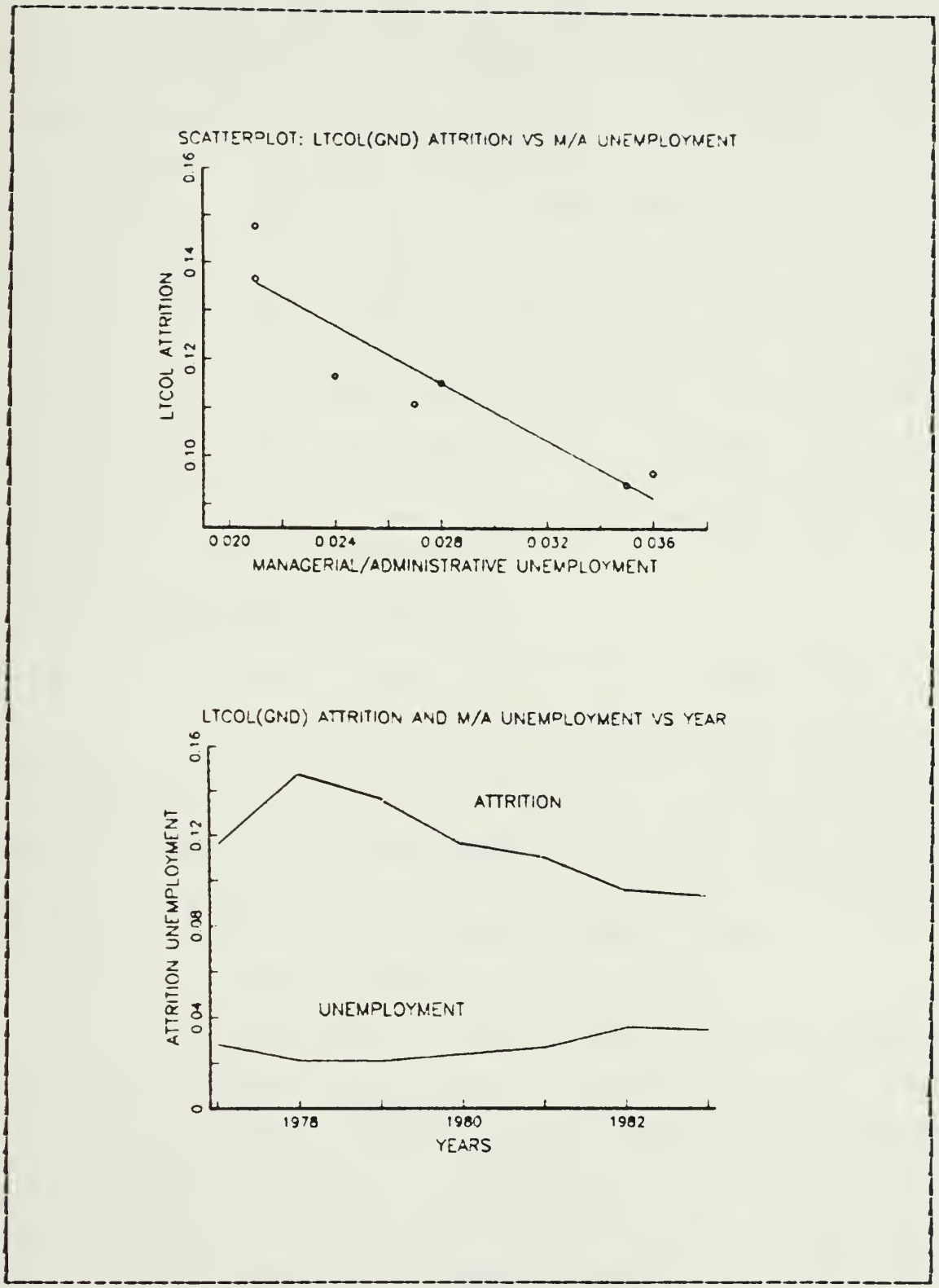


Figure 4.1 Lieutenant Colonel Ground.

A second difficulty was the inability to distinguish between voluntarily separated officers, and those who separated because of their inability to obtain one of the limited number of service extensions. This difficulty was most apparent in the Ground and Total categories of Lieutenants. Because of the great numbers of officers leaving the service in this grade, the effect of the unemployment cycle on attrition is muted and, thus, is not as apparent as in other ranks. Aviation Lieutenants were not included in the model because an analysis of the data disclosed that few (about nineteen per year) ever attrited, and those who did were separated as the result of courts martials or for medical reasons. Based on the above plot, and others not shown, the variables described below were selected for further evaluation.

2. Description of Variables

a. Managerial/Administrative Unemployment (UN)

This is managerial and administrative unemployment and as the name implies, includes executives, managers, and administrators. From this variable, two other lagged variables were created and named UN1 and UN2 (only UN1 is displayed in Table II). The number indicates that variable UN has been lagged one or two years respectively.

b. Professional Technical Unemployment (PT)

This category of unemployment includes lawyers, teachers, computer specialists, airline mechanics, and pilots. It is also computed by the Department of Commerce and is lagged by one and two years in the variables PT1 and PT2 (only PT1 is displayed in Table II). This variable was very significant in the first runs of Captain, Ground and Total. The Department of Commerce, however, has radically

altered this category of unemployment based on the results of the 1980 census to the point where it is no longer consistent with the years prior to 1983 and therefore is unusable in this predictive model.

c. Military/Civilian pay ratio (PAY)

This variable is the ratio of civilian to military pay described in Chapter Three. Its purpose is to display the changing relationship between military and civilian pay through the years. It was also lagged one and two years in the variables PAY1 and PAY2 (only PAY1 is displayed in Table II).

d. Gross Unemployment (GUN)

This is unemployment of males 18 and over. It covers all industries and is one of the rates commonly seen in the newspapers.

3. Promotion Potential

From the discussions with senior officers mentioned in Chapter Three, it became apparent that promotion potential was thought to be a major factor in determining whether or not an officer will remain in the Marine Corps. For example if an officer thought he was promotable to Major and Lieutenant Colonel when he was a Captain there would be a strong likelihood that he would remain in the service. This is true for Lieutenants, Majors, and Lieutenant Colonels as well. A high promotion potential may be indicative of a person's satisfaction with the service. At the same time, not having the possibility for more responsibility because of low promotion potential will cause an officer to leave the service. Thus promotion potential and other factors of retention may be related. The only difficulty was that there were no reliable methods of quantifying this factor, and thus promotion potential was not used in this model.

B. BASIC REGRESSION MODEL

The theory of regression must be understood before the evaluation of the results can begin. The purpose of regression is to validate a theoretical relationship between a given fact and a piece or pieces of information on which it may be dependent. "Running a linear regression" is the procedure by which the coefficients to the variables in the expression will be obtained:

$$Y(i) = a + bX(i).$$

This establishes the linear relation between the X and Y variables. Assumptions made in determining this model are:

(1) There exists a population of Y values for each X; the population random variable corresponding to X(i) is Y(i).

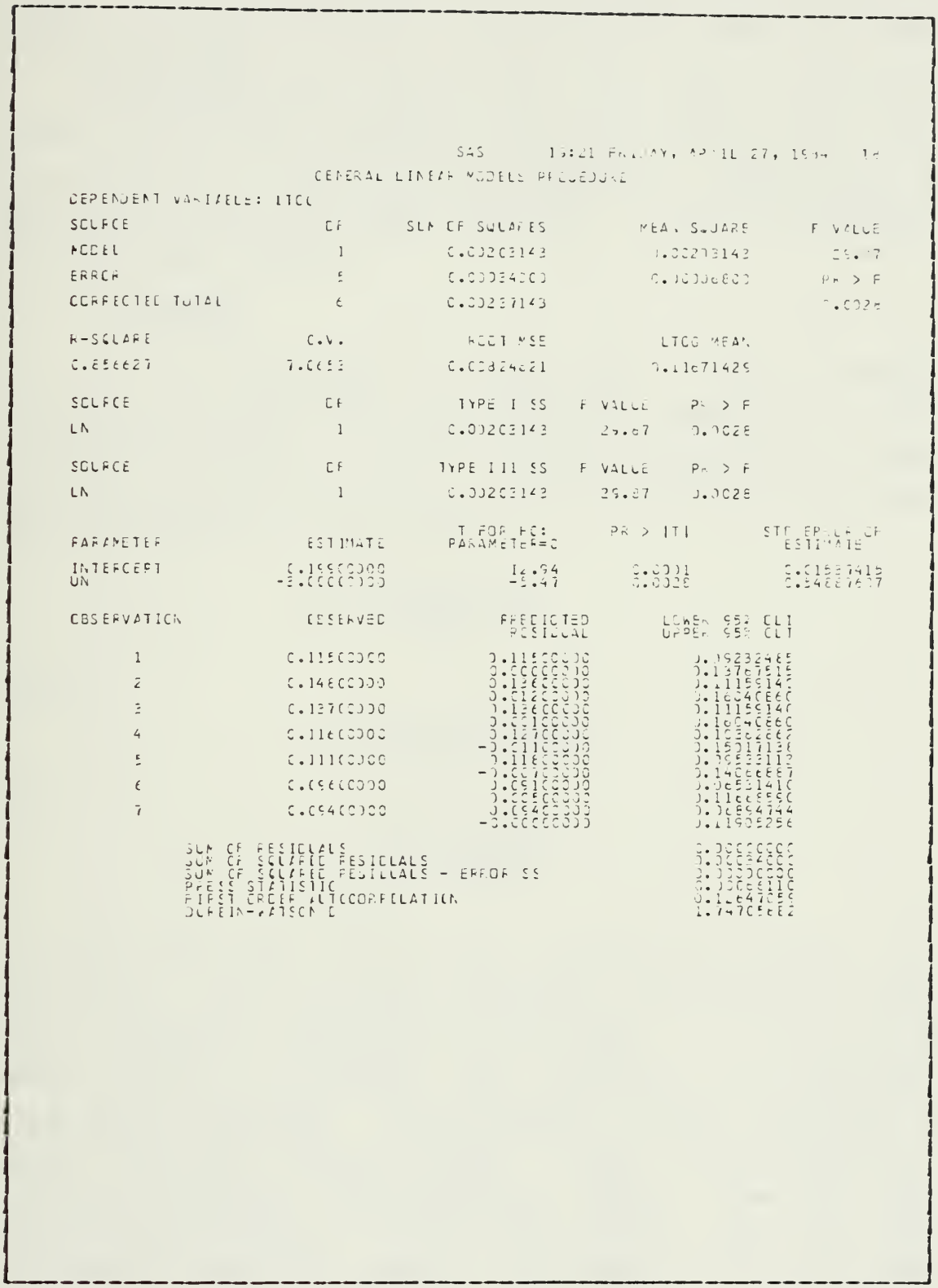
(2) E(Y(i)) = a + bX(i) for each X(i) (i.e., the value of Y that is expected for each X(i) is given by the expression a + bX(i)).

(3) Var(Y(i)) = sigma squared for each X(i).

(4) The errors of observation (residuals) are uncorrelated and normally distributed [Ref. 9].

C. ELEMENTS OF THE SAS OUTPUT

To accomplish the analysis of these variables the SAS Institute's regression analysis programs were used. Figure 4.2 is an example of the output from one of these programs. They provide accurate output with a wide variety of functions. SAS's ease of use and availability were prime factors in its choice for the regression analysis.



```
                                SAS        15:21 FRIDAY, APRIL 27, 1984    14
                        GENERAL LINEAR MODELS PROCEDURE
DEPENDENT VARIABLE: LTCC
SOURCE                      DF       SUM OF SQUARES          MEAN SQUARE       F VALUE
MODEL                        1         0.00203143           0.00203143          25.87
ERROR                        5         0.00039400           0.00007880          PR > F
CORRECTED TOTAL              6         0.00242543                               0.0028

R-SQUARE          C.V.            ROOT MSE              LTCC MEAN
0.837627          7.0653          0.00887621            0.11671429

SOURCE                      DF          TYPE I SS     F VALUE    PR > F
LN                           1         0.00203143       25.87    0.0028

SOURCE                      DF         TYPE III SS     F VALUE    PR > F
LN                           1         0.00203143       25.87    0.0028

                                        T FOR HO:                          STD ERROR OF
PARAMETER                ESTIMATE       PARAMETER=0      PR > |T|          ESTIMATE
INTERCEPT              0.19900000           14.94          0.0001        0.01332415
LN                    -0.06000000           -5.09          0.0028        0.01179697

OBSERVATION            OBSERVED          PREDICTED          LOWER 95% CLI
                                         RESIDUAL           UPPER 95% CLI
        1              0.11500000         0.11500000         0.08232485
                                          0.00000000         0.14767515
        2              0.14800000         0.12600000         0.11159147
                                          0.02200000         0.14040853
        3              0.13700000         0.13600000         0.11159147
                                          0.00100000         0.16040853
        4              0.11600000         0.12700000         0.10259667
                                         -0.01100000         0.15140333
        5              0.11100000         0.11800000         0.09359113
                                         -0.00700000         0.14240887
        6              0.09600000         0.09100000         0.06659141
                                          0.00500000         0.11540859
        7              0.09400000         0.09400000         0.06894744
                                          0.00000000         0.11905256

        SUM OF RESIDUALS                                     0.00000000
        SUM OF SQUARED RESIDUALS                             0.00039400
        SUM OF SQUARED RESIDUALS - ERROR SS                  0.00000000
        PRESS STATISTIC                                      0.00068116
        FIRST ORDER AUTOCORRELATION                          0.12647059
        DURBIN-WATSON D                                      1.74705882
```

Figure 4.2 Example of SAS Output.

Figure 4.3 is an example of the SAS program used to produce the regressions.

In reviewing the results of the SAS output a variety of terms will be used to describe the model and its fit to the data as well as its ability to predict the future based on estimates of the independent variable. A brief description of these terms follows.

1. F value

A statistic used to test for a linear relation between the independent and dependent variables of the regression equation. If the linear relationship is strong we expect that this value, a ratio of explained to unexplained variances, to be large.

$$F = \frac{\hat{b}^2 \sum (Xi-\bar{X})^2}{SE^2}$$

where:

$$b = \frac{\sum (Xi-\bar{X})(Yi-\bar{Y})}{\sum (Xi^2-\bar{X})^2}$$

and:

$$SE = \sqrt{\frac{\sum (Yi-\hat{Y}i)^2}{n-2}}$$

is simply the square root of the unexplained variance. As used in SAS the F statistic measures the explanatory power that a variable contributes to the model. In doing so it tests the null hypotheses that there is no correlation between the variables. If the value of the F statistic is higher than the critical value, which is determined by degrees of freedom and confidence level, then the null hypothesis rejected. If it is lower, then the null hypothesis cannot be rejected. In general, high F values indicate a strong correlation and low values show a weak correlation [Ref. 10].

```
1          S A S     6 0 0     OS SAS 82.3          VS2/4.2 J...  ...  ...  ...  ...
                                                                 13:21 FRIDAY, APRIL 27, 1984

NOTE: THE JOB ESMARTZ HAS BEEN RUN UNDER RELEASE 82.3 OF SAS
      AT NAVAL POSTGRADUATE SCHOOL (6604E0C1).

NOTE: SAS OPTIONS SPECIFIED ARE:
      LINESIZE=80 SOR]=Z

1             DATA;
2                INPUT LN LM1 LTT CAPT MAJT LTCT CCLT;
3                CARDS;

NOTE: DATA SET WORK.DATA1 HAS 7 OBSERVATIONS AND 7 VARIABLES. 317 OBS/TRK.
NOTE: THE DATA STATEMENT USED 0.07 SECONDS AND 208K.

11                ;
12            PROC GLM;
13                MODEL LTT=LN/P CLI;
14                OUTPUT OUT=NEW P=PREDICT R=RESID;
15

NOTE: DATA SET WORK.NEW HAS 7 OBSERVATIONS AND 9 VARIABLES. 250 OBS/TRK.
NOTE: THE PROCEDURE GLM USED 0.29 SECONDS AND 208K
      AND PRINTED PAGE 1.

15                PROC PRINT DATA=NEW;
16

NOTE: THE PROCEDURE PRINT USED 0.20 SECONDS AND 208K
      AND PRINTED PAGE 2.

16            PROC GLM;
17                MODEL CAPT=LN1/P CLI;
18                OUTPUT OUT=NEW P=PREDICT R=RESID;
19

NOTE: DATA SET WORK.NEW HAS 7 OBSERVATIONS AND 11 VARIABLES. 207 OBS/TRK.
NOTE: THE PROCEDURE GLM USED 0.30 SECONDS AND 208K
      AND PRINTED PAGE 3.

19            PROC GLM;
20                MODEL CAPT=LN/P CLI;
21                OUTPUT OUT=NEW P=PREDICT R=RESID;
22

NOTE: DATA SET WORK.NEW HAS 7 OBSERVATIONS AND 10 VARIABLES. 175 OBS/TRK.
NOTE: THE PROCEDURE GLM USED 0.27 SECONDS AND 208K
      AND PRINTED PAGE 4.

22                PROC PRINT DATA=NEW;
23

NOTE: THE PROCEDURE PRINT USED 0.20 SECONDS AND 208K
      AND PRINTED PAGE 5.

23            PROC GLM;
24                MODEL MAJT=LN/P CLI;
25                OUTPUT OUT=NEW P=PREDICT R=RESID;
26

NOTE: DATA SET WORK.NEW HAS 7 OBSERVATIONS AND 12 VARIABLES. 153 OBS/TRK.
NOTE: THE PROCEDURE GLM USED 0.29 SECONDS AND 208K
      AND PRINTED PAGE 6.

26                PROC PRINT DATA=NEW;
27

NOTE: THE PROCEDURE PRINT USED 0.20 SECONDS AND 208K
      AND PRINTED PAGE 7.

27            PROC GLM;
28                MODEL LTCT=LN/P CLI;
29                OUTPUT OUT=NEW P=PREDICT R=RESID;
30

NOTE: DATA SET WORK.NEW HAS 7 OBSERVATIONS AND 17 VARIABLES. 130 OBS/TRK.
NOTE: THE PROCEDURE GLM USED 0.29 SECONDS AND 208K
```

Figure 4.3 Example of a SAS Program.

2. R-square

This statistic is used to examine the linear relationship between the variables of a regression. A perfect relationship would result in a value of 1, while no relationship would produce a value of 0. This statistic can be interpreted as a percent of total variation of the dependent variable that is defined by the independent variable. Normally in time series data with large samples, R-square values of .90 or more are common.

$$\text{R-square} = \frac{\frac{(\hat{Y}i-\bar{Y})^2}{n-2}}{\frac{(Yi-\bar{Y})^2}{n-1}}$$

where Y is the value of Y(i) estimated from the regression equation. [Ref. 11].

3. t-test

This is a test of the hypothesis that the regression coefficients equal zero. The statistic used for this test is the t-statistic:

$$t = \frac{\hat{b}}{S_b}$$

where: b is the estimated coefficient and

$$Sb = \frac{SE}{\sqrt{\Sigma(Xi-\bar{X})^2}}$$

If t is greater than the critical value for the test, then the coefficient b is non-zero. If t is less than the critical value, then the coefficient equals zero, and the null hypothesis that the coefficient has no significance is accepted(Note that for this single variable case t-square equals F) [Ref. 12].

4. Coefficient of Variation

This is the ratio of Standard Error (SE) to the mean of the sample Y's:

$$CV = \frac{SE}{y} \times 100$$

Although the reliability of this expression is subject to the context of its use, values of CV are expected to be as small as .10 to .20 [Ref. 13].

5. Prediction Interval

This is an expression of confidence that a predicted value will be within a pair of values. Thus it can be said, for example, that the attrition rate for 1984 will be within .12 and .13 with 95 percent confidence [Ref. 14].

6. Autocorrelation

Dependence of the value of one variable on the values of the same variables preceding it in time. For example, the dependence of 1984 unemployment on 1983 unemployment would be first order autocorrelation. The degree of autocorrelation is measured by the autocorrelation coefficient rho, p. If p=0 there is no autocorrelation. For positive or negative autocorrelation the values of p are positive or negative. In this paper the term serial correlation is used interchangeably with autocorrelation [Ref. 15].

7. Durbin-Watson Test

This test measures the degree autocorrelation. Comparison with tables will indicate whether the statistic is significant or not. Values of 2.00 indicate no serial correlation. For purposes of this paper, existing tables

had to be extended by linear extrapolation to cover small sample sizes. The resulting tests are as follows:

For negative autocorrelation:

Ho: No negative autocorrelation
Reject if DW < 4-dl
Accept if du < DW < 4-du
Inconclusive if 4-du < DW < 4-dl

For positive autocorrelation:

Ho: No positive auto correlation
Reject if DW < dl
Accept if DW > du
Inconclusive if dl < DW < du

Values of du and dl for this problem are 1.28 and .92 respectively [Ref. 16].

D. EFFECTS OF SAMPLE SIZE ON REGRESSION

As previously mentioned, in the construction of this model there were two significant problem areas. Both problems are a result of the lack of data points. In performing linear regression, the larger the sample size the greater is the possibility that the relationship indicated by the model is a statistically valid one. An R-square of .90 with one thousand data points is far more likely to be a valid model than a model with the same R-square and only seven data points. In a time series model, for instance, a relationship that fits the data well over fifty years is far more likely to be valid than one that fits over ten years. Thus, this model with only seven data points has obvious questions of validity which only additional data will be able to confirm.

The second consequence of the small sample sizes is the restriction that must be placed on introducing additional

variables to the model. It is unlikely that the variable UN explains all of the variation in the attrition rate. Intuitively, there must be other factors. Unfortunately, with a small sample size, there is a tendency for model statistics such as the F value and R-square to improve merely by the addition of a variable or two without these variables also having a significant effect on the model's description of reality. For instance, R-square will increase to one if six variables are added to a model with a sample size of six. This is a result of the method by which R-square is calculated. Thus to ensure that an accurate model is produced, the number of independent variables has been limited to one, despite the fact that there are other variables that have significance both intuitively and statistically and, in general, the model is improved by adding them.

E. DIFFICULTIES WITH AUTOCORRELATION

The problem of autocorrelation was encountered in four of the models. This is normal in time series data and in large data sets there are standard procedures that attempt to remove the autocorrelation. In the case of small data sets, however, the procedures are sometimes not effective and produce results which are subject to question. The presence of autocorrelation does nothing to hamper the model's predictive consistency, but it does cause the estimate of standard error to be biased [Ref. 17]. The goal of any procedure intended to correct autocorrelation is to produce residuals which are uncorrelated and thus satisfy the independence assumption of least squares linear regression. The results of these models with their evaluatory statistics are shown below in Table III.

The procedure used involved estimating the serial correlation coefficient using the Hildreth-Lu procedure. The Hildreth-Lu procedure is a grid search method of determining the optimal serial correlation coefficient by computing the sum of squared residuals for a series of possible values of the serial correlation coefficient. Values of the grid range from 1.0 to -1.0 in gradations of .1, including zero. The optimal value of the correlation coefficient is then the one that produces the minimum value of the residual sum of squares resulting from the regression using the general differencing procedure (see below) for each correlation coefficient. The optimal value of the correlation coefficient to the one-hundredths place was then determined by a second grid search covering an equal distance to either side of the initial point. After determining the optimal estimate of the serial correlation coefficient, the general differencing procedure [Ref. 18] was used to apply this factor to the regression model in the form:

$$Y^* = Bo(1-p) - B1(Xt^*)$$

where:

$$Y^* = Yt - p(Yt-1)$$
$$X^* = Xt - p(Xt-1)$$

and p = serial correlation coefficient
Xt-1 = previous year's unemployment
Xt = follow-on year's unemployment
Y* = attrition rate prediction for follow-on year

The results of this procedure are discussed in the following section and are presented with their statistics in Table III. Comparisons of the equations arrived at by general differencing versus the original equations are found in Appendix B. In three of the four cases the model produced by the general differencing was superior to the

original model both in the Durbin-Watson statistic as well as in the other measures of a model's validity (it is noted that after general differencing the normal distribution theory for linear models no longer holds).

F. VERIFICATION OF REGRESSION ASSUMPTIONS

After each model was selected, the residuals were examined for normality and independence. Despite the fact that these or any procedures are relatively inaccurate for small sample sizes, normal plots were used together with the Shapiro-Wilk statistic to verify normality. The results of these tests generally indicated normality. Those results (3) that were inconclusive or weak involved the models on which general differencing was used. The use of this procedure caused the loss of one more data point and thus the probable weakening of the tests. Constant variance of the residuals was established by the examination of plots of the residuals against the estimated attrition and the independent variable, while independence was established by the use of the DW statistic and graphical methods.

G. MODEL VARIABLES AND COEFFICIENTS

Table IV is a tabular presentation of each model's independent variable with it's coefficients. Where a value for Rho appears in the right hand column general differencing was used to achieve the final model. The statistical quality of each model is discussed in detail below.

H. APPLICATION OF STATISTICAL TESTS TO THE MODEL

The following is a grade-by-grade analysis of the model based on the test statistics provided by the SAS output. The following discussion will refer to Table V concerning

TABLE III

RESULTS OF GENERAL DIFFERENCING PROCEDURE

Rank/comp	Var (1)	R (2)	F (3)	ta (4)	tb (5)	DW (6)	CV (7)
Major							
Ground	un	.97	130	23.4	-11.4	2.78	.04
Major							
Aviation	un	.73	10.9	6.0	-3.3	2.33	.27
Captain							
Total	un	.99	281	45.3	-16.8	2.96	.02
Lieutenant							
Ground	un	.53	4.5	5.7	-2.1	1.70	.32

(1) The independent variable used in the regression.
(2) R-squared value
(3) F statistic
(4)-(5) t statistic for coefficients a and b
(6) Durbin-Watson statistic
(7) Coefficient of Variance

the statistical tests of the model. The results of the comparison of the general differencing model and the original model are included in Table V. Columns are identified and explained beneath the table.

TABLE IV

MODEL VARIABLES AND COEFFICIENTS

Rank/comp	Variable	a-coeff	b-coeff	Rho
Lt Col				
Aviation	UN	.238	-4.078	-
Ground	UN	.199	-3.000	-
Total	UN	.208	-3.194	-
Major				
Aviation	UN	.049	-1.784	.49
Ground	UN	.251	-2.677	-.67
Total	UN	.140	-2.637	-
Captain				
Aviation	UN	.294	-5.625	-
Ground	UN1	.133	-1.678	-
Total	UN	.136	-2.394	.23
Lieutenant				
Ground	UN	.0597	-2.342	.72

1. Lieutenant Colonels

a. Aviation.

The aviation model for Lieutenant Colonel was negatively affected by the unnatural retention of officers in this grade resulting from the enactment of Aviation Officers Continuation pay. The attrition rate of aviation Lieutenant Colonels dropped more than fifty percent from the year before. With the modification of this value to reflect the absence of those officers electing the continuation pay,

TABLE V

STATISTICAL ANALYSIS OF THE MODEL

Rank/comp	Var (1)	R (2)	F (3)	ta (4)	tb (5)	DW (6)	CV (7)
Total							
Aviation	UN	.54	5.8	5.0	-2.4	2.06	.20
Ground	UN	.86	30	13	-5.5	1.75	.07
Total	UN	.78	18	10	-4.2	2.06	.09
Major							
Aviation	UN	.73	11	6.0	-3.3	2.33	.27
Ground	UN	.97	131	23	-11	2.78	.04
Total	UN	.88	36	11	-6.0	1.39	.10
Captain							
Aviation	UN	.98	302	33	-17	2.00	.03
Ground	UN1	.76	16	12	-4.0	2.31	.06
Total	UN	.88	36	16	-6.0	.90	.05
Lieutenant							
Ground	UN	.56	6.2	6.0	-2.5	.65	.15

(1) The independent variable used in the regression.
(2) R-squared value
(3) F statistic
(4)-(5) t statistic for coefficients a and b
(6) Durbin-Watson statistic
(7) Coefficient of Variance

who otherwise would have attrited, the model would improve immensely. Unfortunately these officers are impossible to identify. As more data is developed the inclusion of a pay variable in this model may increase its efficiency. An

additional consideration in explaining the drop in attrition is the simple fact that Lieutenant Colonels are the rank group most severely affected by a poor economy if they leave the service. This is because the typical Lieutenant Colonel is married with two college-age children and a large home mortgage. The model is significant at the .93 level and the Durbin-Watson statistic of 2.06 indicates that there is no serial correlation present in the residuals since 2.06 > 4-dl and 2.06>du for negative and positive serial correlation, respectively.

b. Ground.

As can be seen in the table, the statistics evaluating this model indicate that it is an accurate model despite the paucity of data. The model is significant at the .99 level as are the a and b coefficients. The coefficient of variation is less than .1 and the Durbin-Watson statistic indicates that there is no significant autocorrelation present.

c. Total.

This model is a composite of aviation and ground components and the statistics react accordingly. All worsen except the Durbin-Watson statistic which at 2.06 indicates no significant serial correlation.

2. Major

a. Aviation.

The same difficulties encountered in modeling aviation Lieutenant Colonels were encountered with Majors. Again, a large proportion of Majors who would normally attrite accepted Aviation Officer Continuation Pay (AOCP) and did not leave the service in 1982 and 1983.

Additionally, Majors who reached retirement eligibility because of prior enlisted service chose to wait until their mandatory retirement point at twenty years of commissioned service because of the poor state of the economy. The F statistic is significant at the .95 level as are the t statistics for both coefficients. The coefficient of variation is fairly good at .21 compared to the .10 to .20 normally desired. This value is consistent with the R-square of .60. In this case a pay variable that took into consideration the effect of AOCP might be a valid second independent variable to be included as the size of the data set increases. The Durbin-Watson test indicates that there is positive serial correlation present in the residuals. The general differencing procedure was used on this model, and the result was a model which was vastly superior statistically with no serial correlation.

b. Ground.

The relation between ground Major attrition and unemployment is quite strong with an R-square of .82 and a coefficient of variation of .11. Both the a and b coefficients are significant at the .99 level and the F statistic at 22.9 is also significant at the .99 level. The Durbin-Watson test is inconclusive suggesting that the correlation of the errors is due to the autocorrelation of the independent variable and not the correlation of residual terms. The result of using general differencing on this model was to greatly improve its evaluatory statistics. The DW statistic, however, only improved from 2.89 to 2.78. This was enough, however, to accept the hypothesis of no negative or positive serial correlation in the model.

c. Total.

This model is better than its individual aviation and ground components. Its R-square is high at .88 and the F and t values for both coefficients are significant at the .99 level. With a Durbin-Watson value of 1.39, the null hypothesis of no serial correlation is not rejected. Also meeting acceptable standards is the low value of CV at .10.

3. Captain

a. Aviation.

Although attrition rates dropped by almost forty percent in 1982 and 1983 from 1981, the model still fits the data extremely well since the reduction in attrition due to AOCP occurred at the same time as the last rise in unemployment rates. A check of Table V shows that all statistics are significant at the .99 level and the Durbin-Watson statistic is 2.00, indicating the absence of serial correlation.

b. Ground.

The R-square for ground captains indicates a relatively good fit of the data. The F statistic is significant at the .99 level as are the t statistics for each coefficient. The coefficient of variation is exceptionally good at .06. The Durbin-Watson statistic of 2.50 indicates the absence of significant serial correlation since du < 2.50 < 4-du.

c. Total.

Based on the relevant statistics the original model was effective except that the Durbin-Watson statistic of .90 is slightly less then the dl of .92 indicating the possible presence of positive serial correlation. The DW

statistic resulting from the general differencing was inconclusive as to the presence of positive or negative serial correlation although the model itself was vastly improved over the original.

4. Lieutenants

As described in previous chapters, this grade is difficult to model since only a limited number of officers are allowed to remain in service. Thus, effects due to unemployment are masked by the large number of officers attriting simply because there is no requirement for them in the Marine Corps. Correspondingly, this model is poor in its explanation of the variance of the data measured by R-square. The F statistic was significant at the .94 level as was the t statistic for the beta coefficient while the alpha coefficient was significant to the .99 level. These statistics were also subject to question because of the presence of positive serial correlation indicated by Durbin-Watson statistic of .70. General differencing produced a slightly worse model with a lower R-square and less significant F and t statistics (.90 level). Serial correlation, however, was eliminated as evidenced by the DW statistic value of 1.70.

V. RESULTS AND CONCLUSIONS

A. MODEL COMPARISONS

Previous efforts by Headquarters Marine Corps (HQMC) had used an averaging process to compute officer attrition estimates needed for planning purposes in the following year. The results of this process, which was merely a sum of previous six years of attrition divided by six, was then modified up or down depending on the intuition of the responsible officer regarding trends in officer retention. For example, a severe recession might result in the total being reduced by an arbitrary percentage if the responsible officer was of the opinion that economic factors would slow down attrition.

As can be imagined, this method was inaccurate in the sense that it would lag the current rate because of the dependence on previous years' rates. Additionally it would not be able to predict the critical turning points in the trend of officer attrition as well as not being able to indicate extremes. Table VI presents a comparison of the averaging method with the results of the linear regression models developed in this research for the year FY 1983 and FY 1984. The third column contains the actual attrition figures for 1983. The fourth and fifth columns show the percent error for the average and 1983 regression model's predictions for 1983 respectively. The sixth column gives the 1984 regression model predictions for 1984. The disparity between the averaging method and what actually happened in 1983 is a result of trends in attrition which averaging cannot predict. These trends are caused by a variety of factors such as the economy, military pay, and military policies such as flight bonuses.

TABLE VI

COMPARISON OF MODEL RESULTS

Rank	Avg Method 1983	Regress Model 1983(1)	Actual 1983	% Error Avg 1983	% Error Reg 1983	Regress Model 1984(2)
LTCOL	177	152	127	39.4	19.7	152
MAJOR	186	149	146	27.4	2.1	164
CAPT	452	377	392	15.3	-3.8	407
1ST LT	525	416	396	32.6	5.1	425(3)

(1) Uses actual value for UN (.035) in 1983. Both the 1983 and 1983 models use 1982 average strength.

(2) Assumes UN=.033 and updated coefficients for 1983.

(3) Includes average annual aviation Lieutenant attrition of 20 per year.

The complete attrition model is presented in table IV of Chapter Four and is in the form of eleven regression relations and includes the modifications resulting from general differencing. The equations were maintained in this form for simplicity of use by HQMC.

E. PREDICTION INTERVALS

Prediction intervals provide a zone of confidence within which it can be claimed that a result will lie with 95 percent probability [Ref. 19]. Prediction intervals are shown graphically by the solid lines on either side of the straight regression line. The graph shown is for Total Lieutenant Colonel attrition versus the unemployment variable UN. All other prediction interval graphs appear in Appendix B.

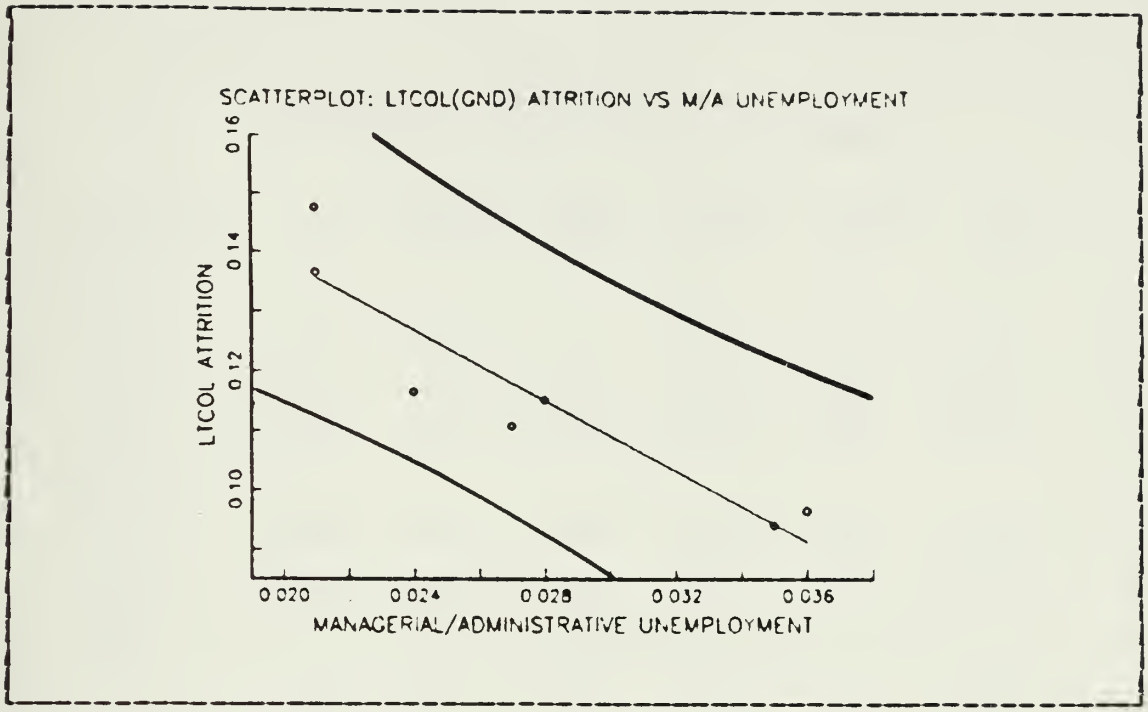


Figure 5.1 Prediction Intervals.

C. SENSITIVITY ANALYSIS

Sensitivity analysis is the procedure by which the model is tested to determine its reaction to various input values. Table VII reflects the general trend of the economy for 1984 with most of the values for unemployment (UN) decreasing from 1983's .035. Also shown in the first two columns are the results for 1984 if unemployment increases. The model attrition rates are relatively insensitive to change as a .001 change in Unemployment produces only five more Lieutenant Colonel attritions. Likewise a change from .035 to .032 in unemployment results in an additional fourteen attritions. (This data was calculated by multiplying 1984 estimated Lieutenant Colonel strength by the change in the attrition rate.) This is more significant and represents an eleven percent increase from FY 1983.

TABLE VII

SENSITIVITY ANALYSIS OF 1984 MODEL

Rank/Comp	Possible Values of UN for 1984 .037	.036	.035	.034	.033	.032
Total						
Itcol	.0897	.0929	.0961	.0993	.1025	.1057
Major	.0424	.0451	.0477	.0504	.0530	.0556
Capt	.0892	.0910	.0929	.0947	.0966	.0984
Ground						
Itcol	.0880	.0910	.0940	.0970	.1000	.1030
Major	.0440	.0485	.0530	.0575	.0619	.0664
Capt	----------	----------	.0731 (1)	----------	----------	----------
It	.1384	.1391	.1397	.1404	.1410	.1417
Aviation						
Itcol	.0836	.0880	.0924	.0967	.1011	.1055
Major	.0332	.0341	.0350	.0359	.0368	.0377
Capt	.0880	.0910	.0940	.0970	.1000	.1030

(1) Since Captain ground depends on the lagged variable UN1 (.035) there is only one value predicted for 1984.

D. CONCLUSION

Based on the results enumerated in the foregoing paragraphs it can be concluded that economic factors were significant variables in the attrition of Marine Officers during the period 1977 to 1983. The economic factor most important in determining attrition is unemployment. A second significant factor may be pay. As more attrition data is compiled these conclusions may well be strengthened if, as predicted, the economy continues to strengthen, and

initial reports of increased officer attrition continues to hold true through the remainder of FY 1984. These trends can also be affected by management policy such as the authorization of bonus payments as in the case of AOCP for Marine Aviators. As explained in Chapter Four, AOCP, in conjunction with a crippled economy, had a significant effect on aviation attrition, reducing the totals by almost forty percent. An attempt was made to develop quarterly models and by this method have more data points available, but the models had little correlation with officer attrition. This was intuitively understandable since officer attrition will almost always occur around the summer months for a variety of personal factors. This is not a result of unemployment but simply a function of when school vacations begin and officers are normally commissioned. Recommendations for further study center around promotion potential coupled with unemployment as a significant indicator of an officer's intentions at the breakpoints of his career in a binary choice model.

APPENDIX A

SCATTER PLOTS AND COMPARISON DIAGRAMS

The following graphs are the scatter plots and comparison diagrams for all rank/components included in the model. Many other plots of these types were done for all of the variables considered to include rates of change of the variable but are not included because of the lack of space.

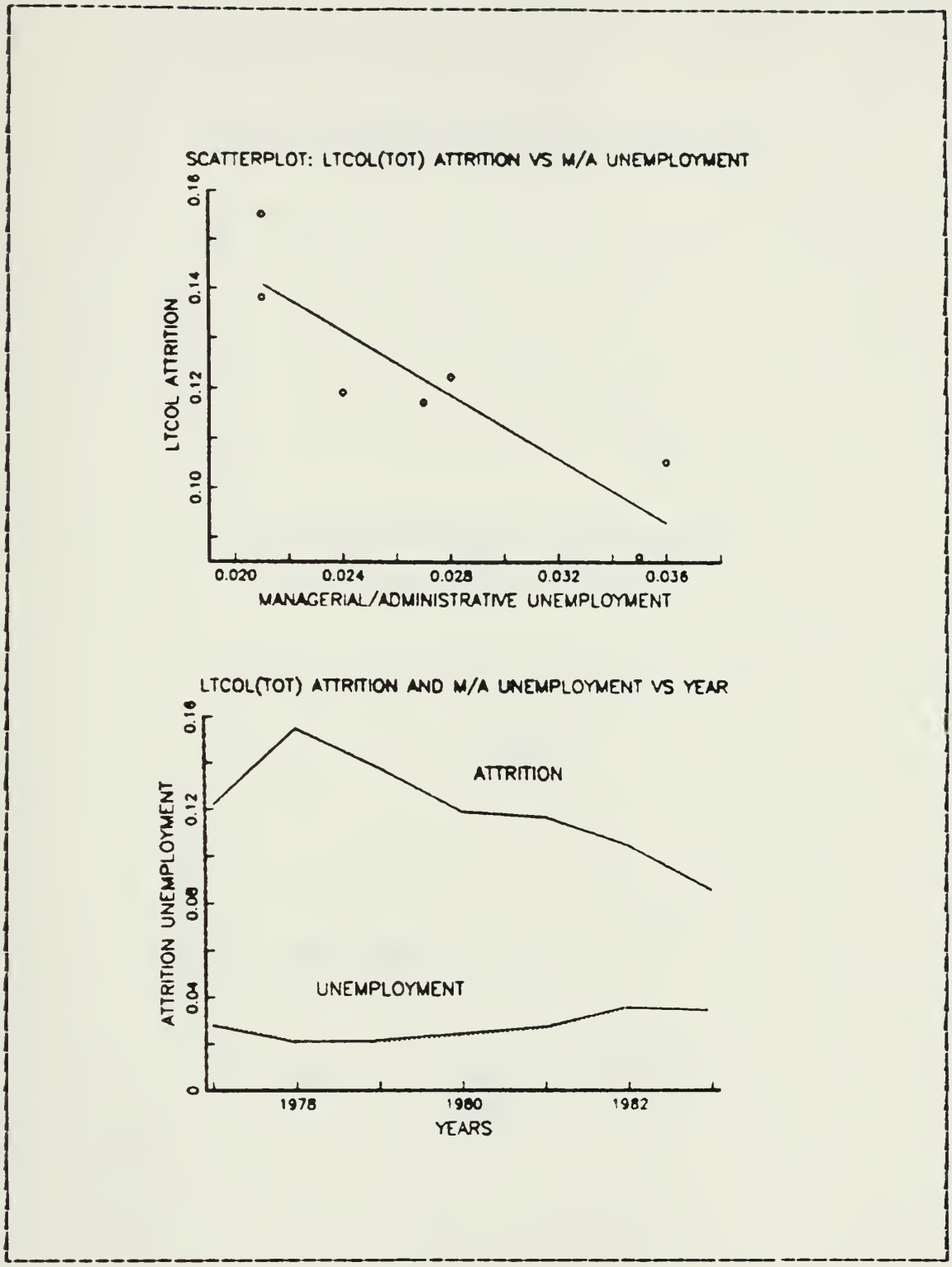
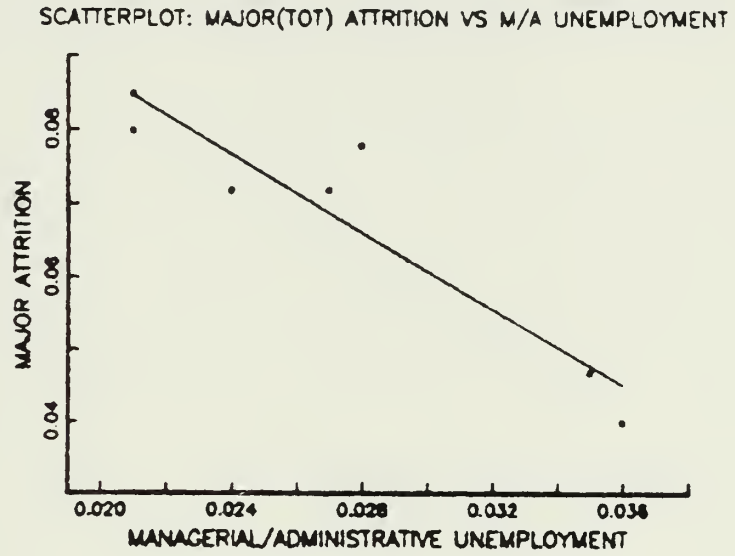


Figure A.1 Total Lieutenant Colonels.



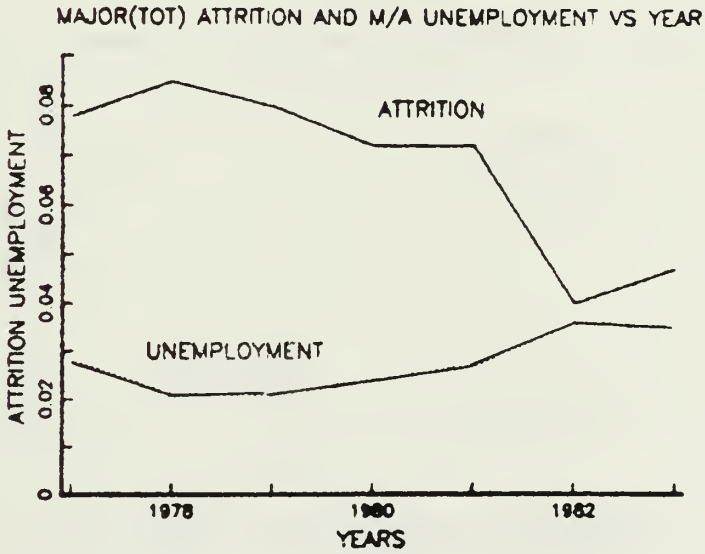
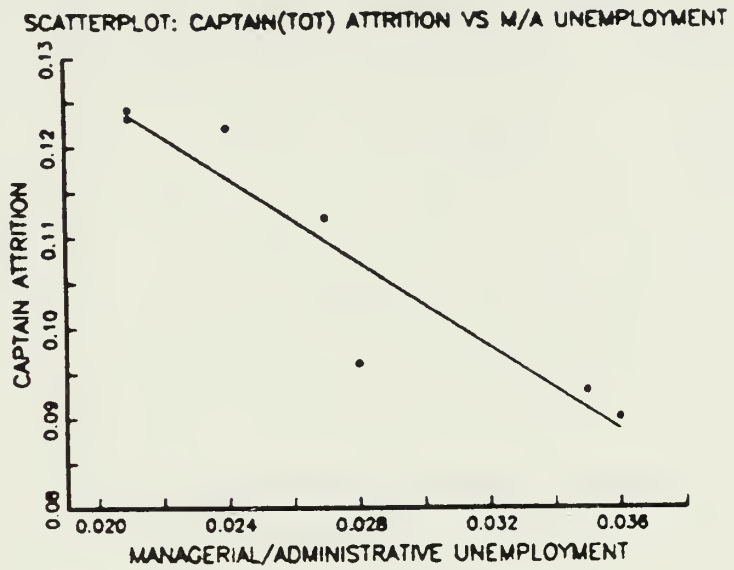


Figure A.2 Total Majors.



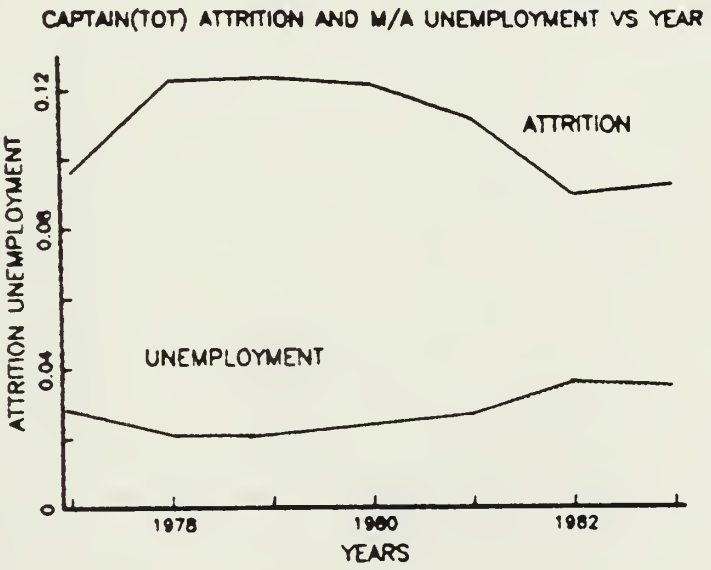
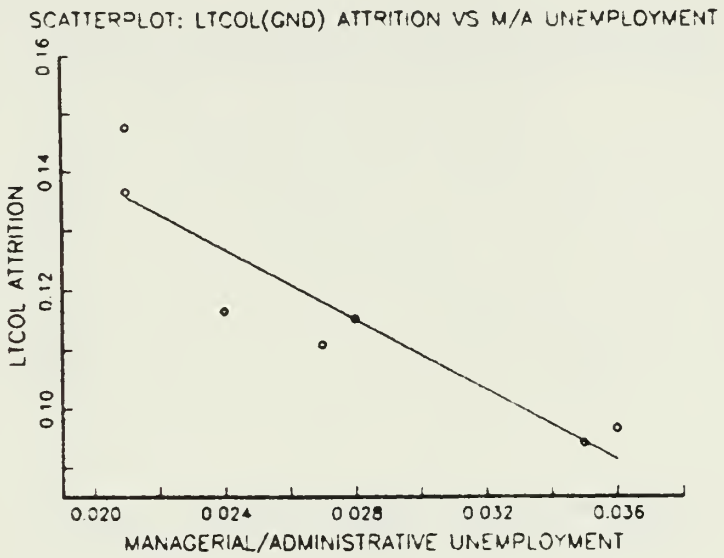


Figure A.3 Total Captains.



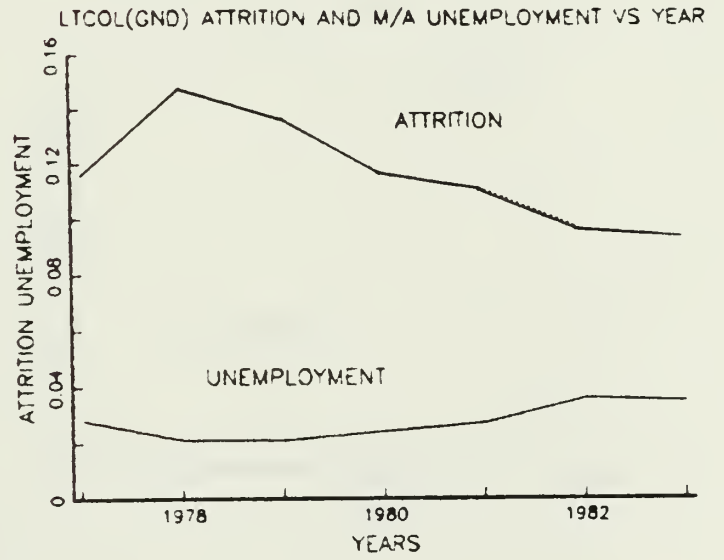
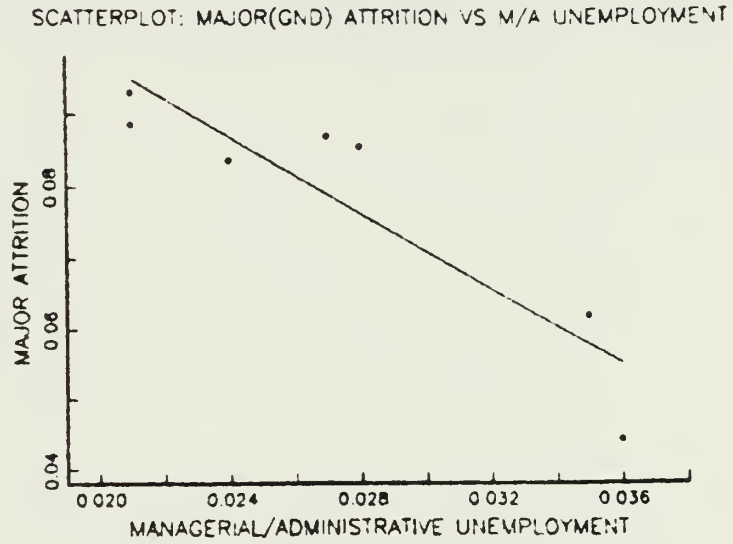


Figure A.4 Ground lieutenant Colonels.



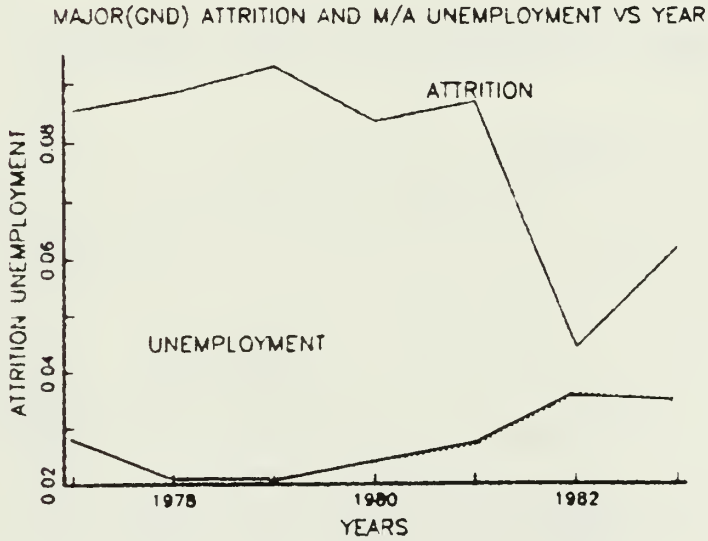


Figure A.5 Ground Majors.

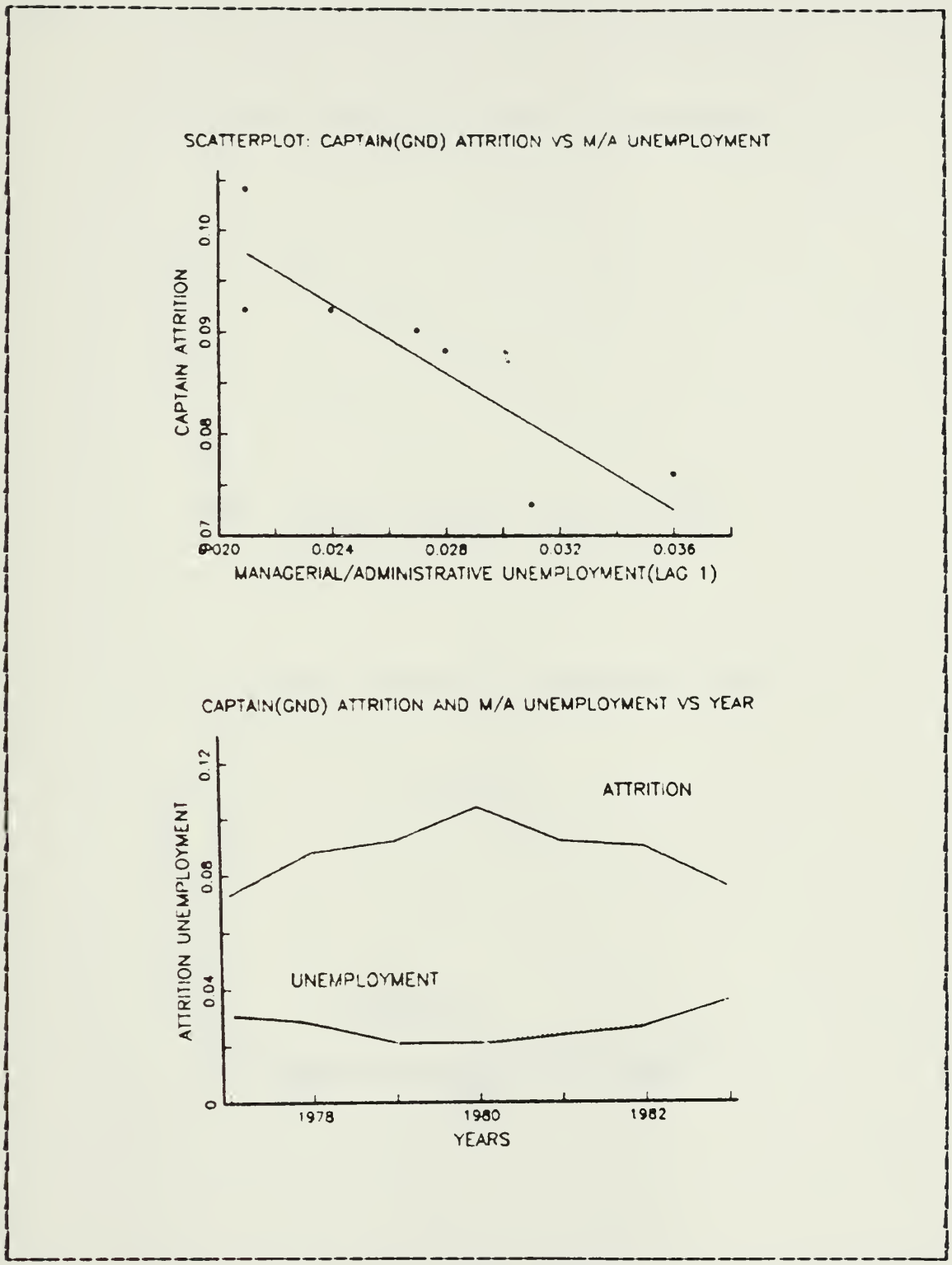


Figure A.6 Ground Captains.

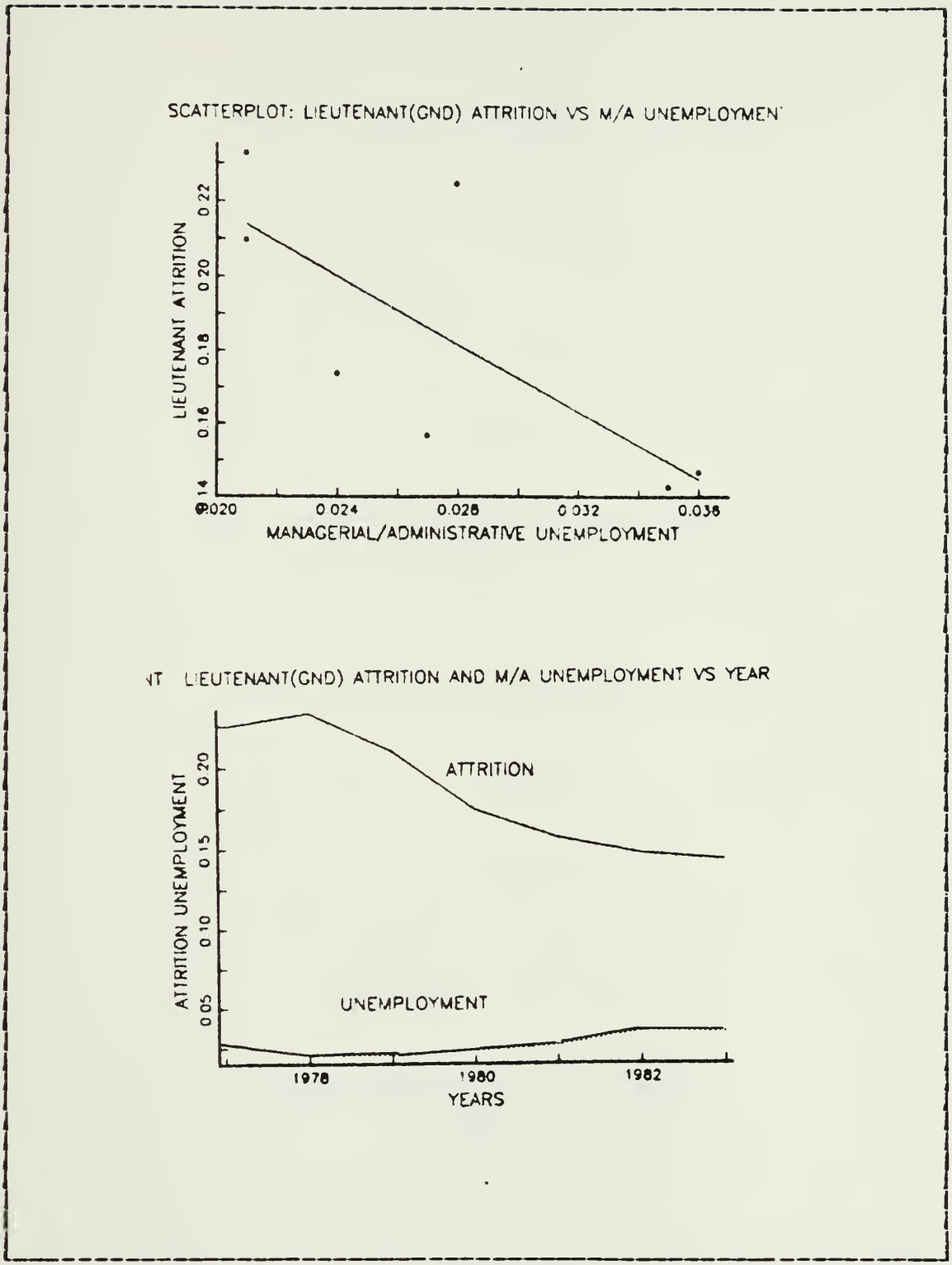


Figure A.7 Ground Lieutenants.

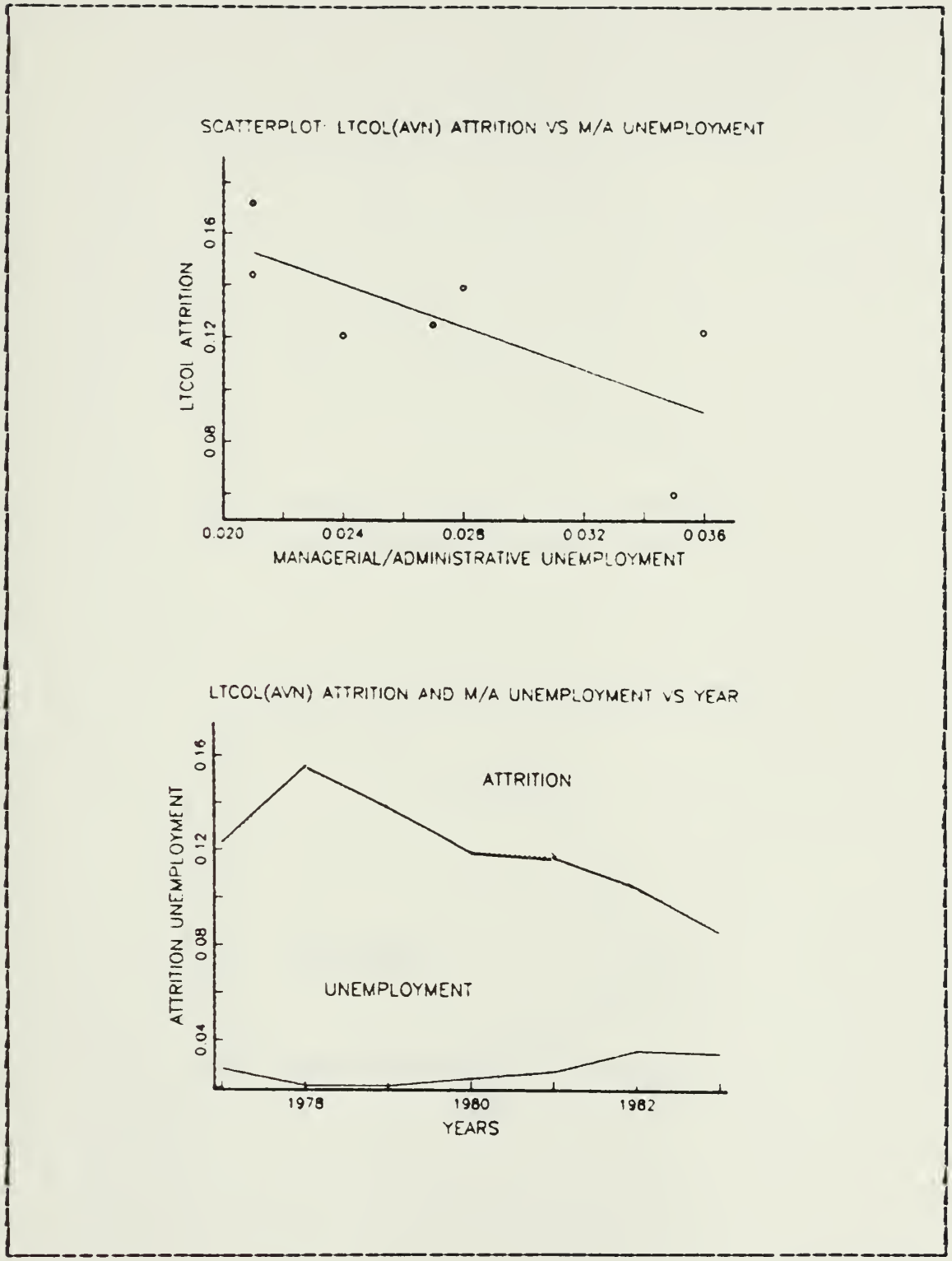


Figure A.8 Aviation Lieutenant Colonels.

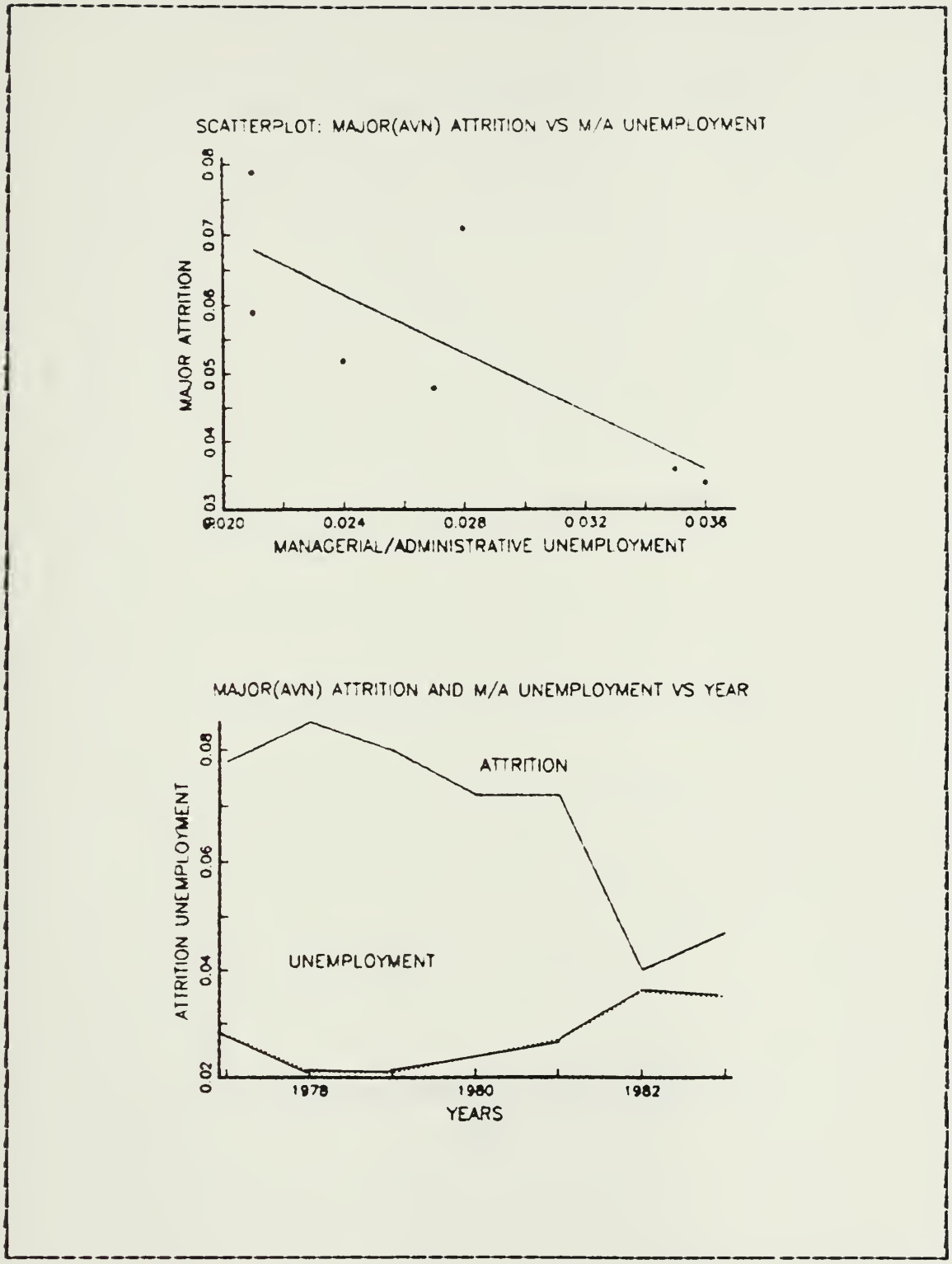


Figure A.9 Aviation Majors.

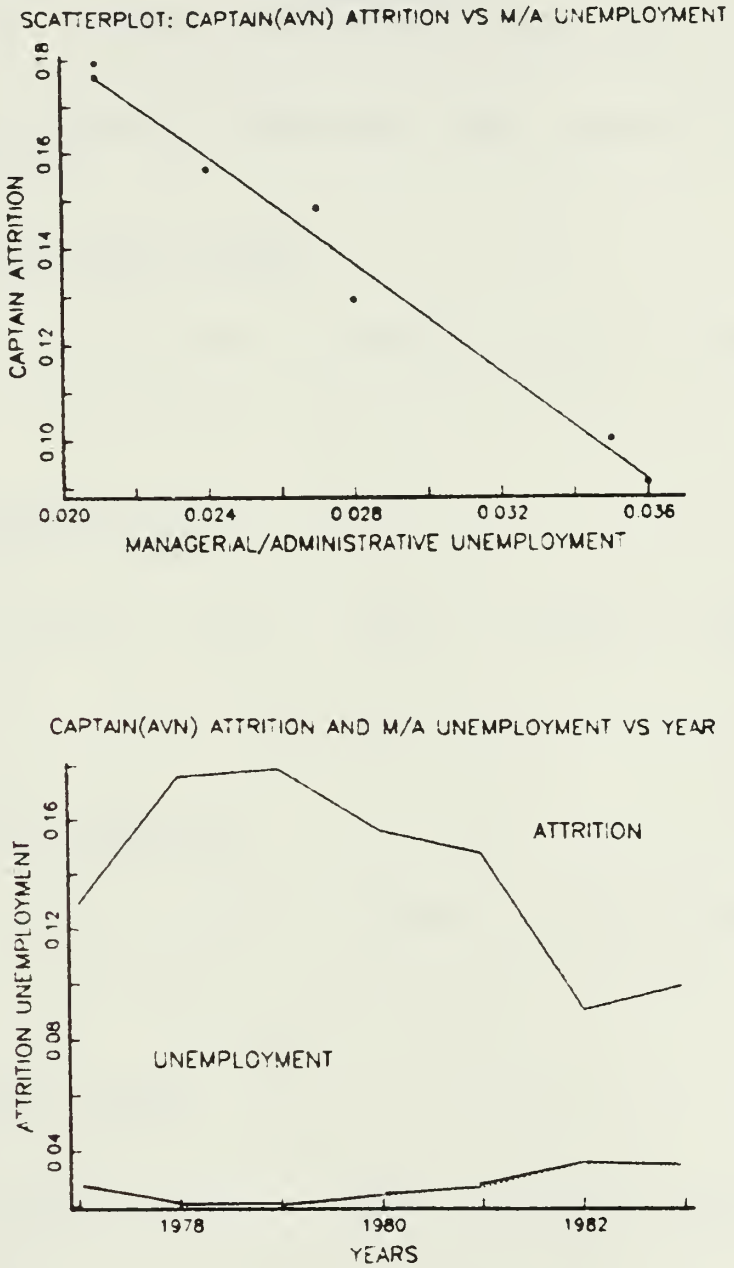


Figure A.10 Aviation Captains.

APPENDIX B
GENERAL DIFFERENCING EQUATIONS

A. MODEL EQUATIONS RESULTING FROM GENERAL DIFFERENCING PROCEDURES.

1. Lieutenant ground.

y = (.143 x .72) + .0597 - 2.342 x (UNt - (.72 x UNt-1))

2. Captain total.

y = (.093 x .23) + .136 - 2.3941 x (UNt - (.23 x UNt-1))

3. Major aviation.

y = (.036 x .49) + .0492 - 1.784 x (UNt - (.49 x UNt-1))

4. Major ground.

y = (.062 x -.67) + .251 - 2.677 x (UNt - (-.67 x UNt-1))

E. COMPARISON OF ORIGINAL

In the following table the original models are composed with the models resulting from general differencing. Model 1 statistics are from the original Model while Model 2 indicates the general differencing model. As can be seen all models improve except the Lieutenant ground model.

TABLE VIII

RESULTS OF GENERAL DIFFERENCING

Rank/comp	Model	R	F	ta	tb	DW	CV
Lt(gnd)	1	.56	6.2	6.0	-2.5	.65	.15
	2	.53	4.5	5.6	-2.1	1.70	.32
Capt(tot)	1	.88	36	16	-6.0	.90	.05
	2	.99	281	43	-17	2.60	.02
Major(av)	1	.60	7.6	5.2	-2.8	.71	.21
	2	.73	11	6.0	-3.3	2.33	.27
Majcr(gnd)	1	.82	23	9.7	-4.8	2.89	.11
	2	.97	131	23	-11	2.78	.04

APPENDIX C
PREDICTION INTERVALS

The following figures show the prediction interval diagrams for the models of all ranks and components. As was expected from the model data some of the models are poor regarding their variation in predicted rates. On these models where the prediction intervals are too large to be shown without expanding the scale they are not shown. These models in which general differencing was used follow the original model.

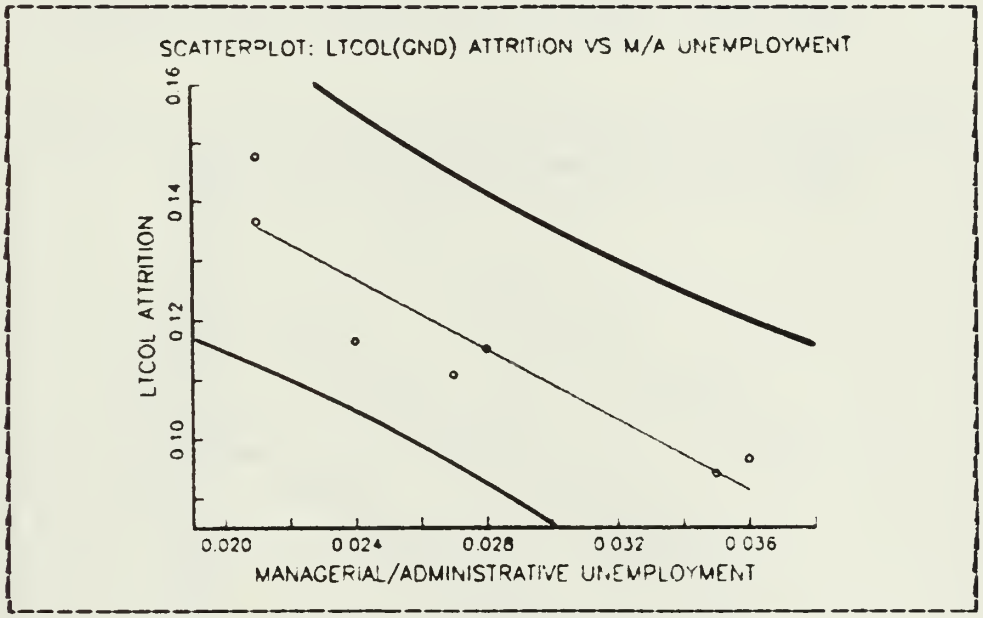


Figure C.1 Lieutenant Colonel Ground.

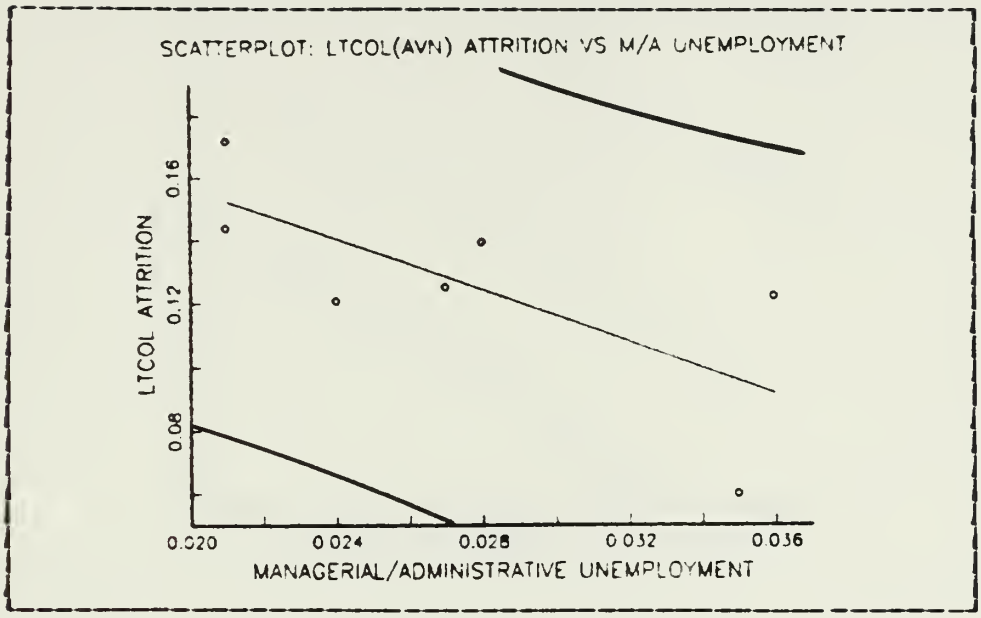


Figure C.2 Lieutenant Colonel Aviation.

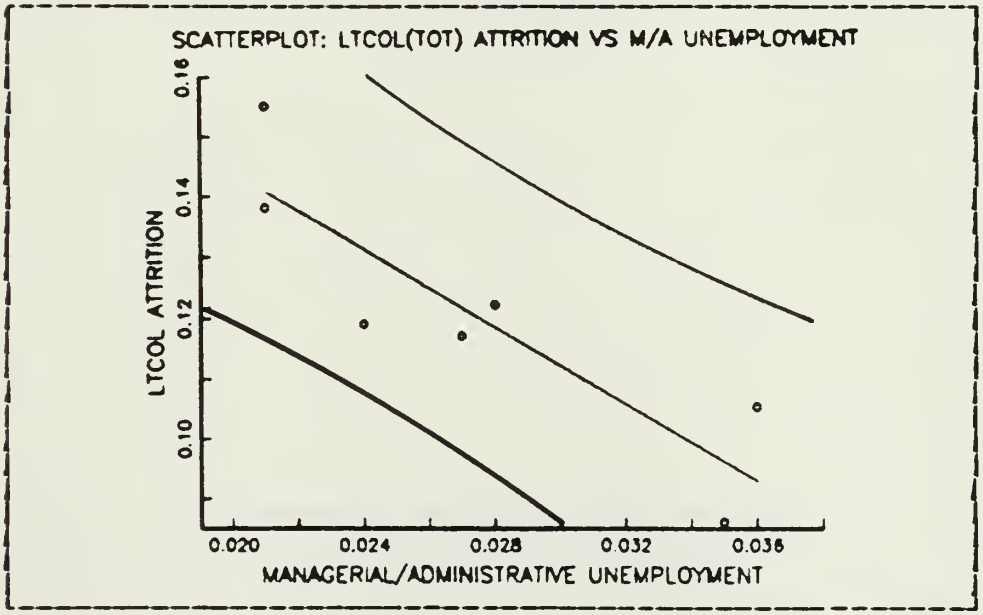


Figure C.3 Lieutenant Colonel Total.

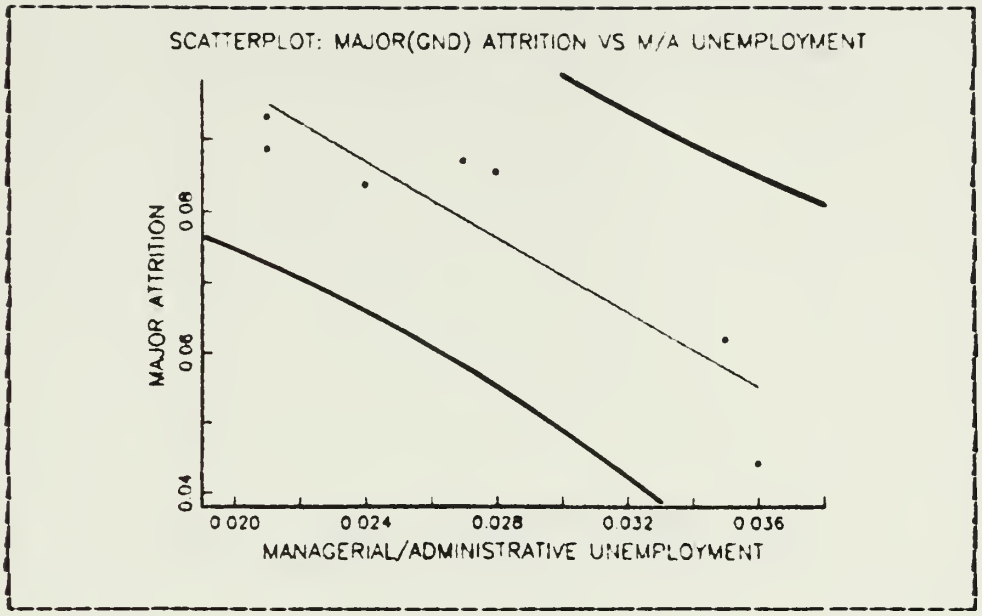


Figure C.4 Major Ground.

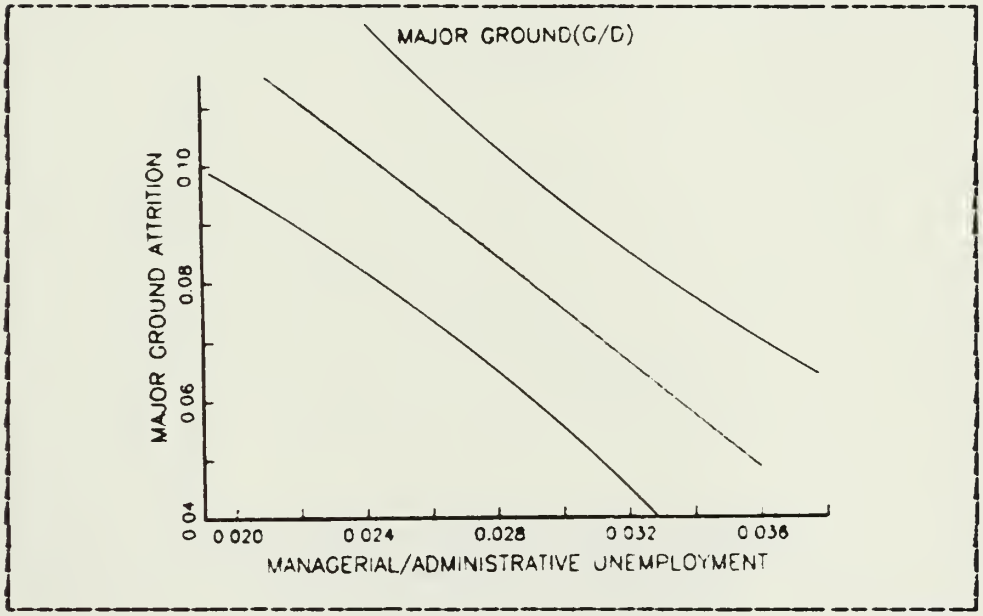


Figure C.5 Major Ground/Using General Differencing.

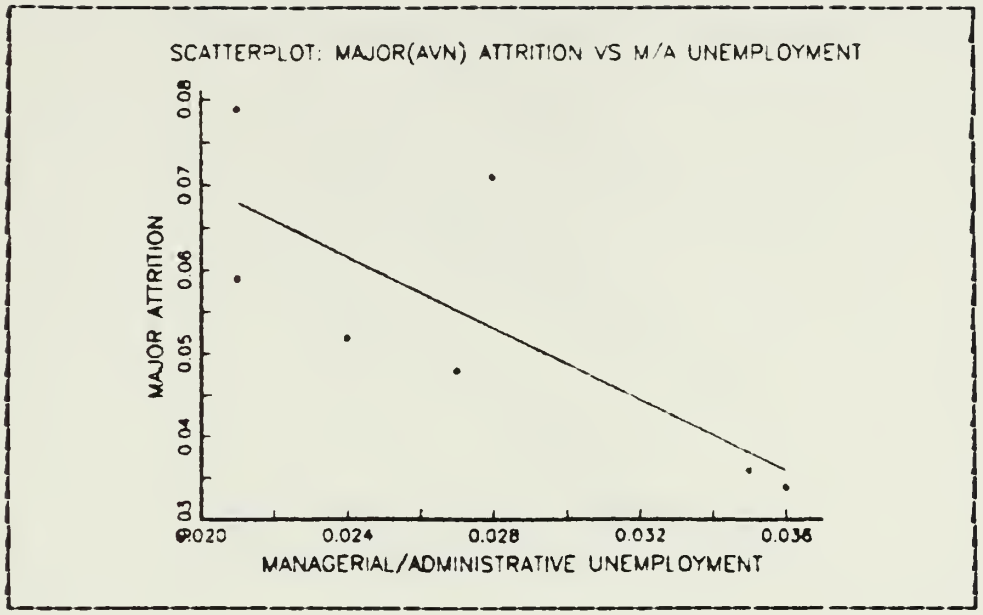


Figure C.6 Major Aviation.

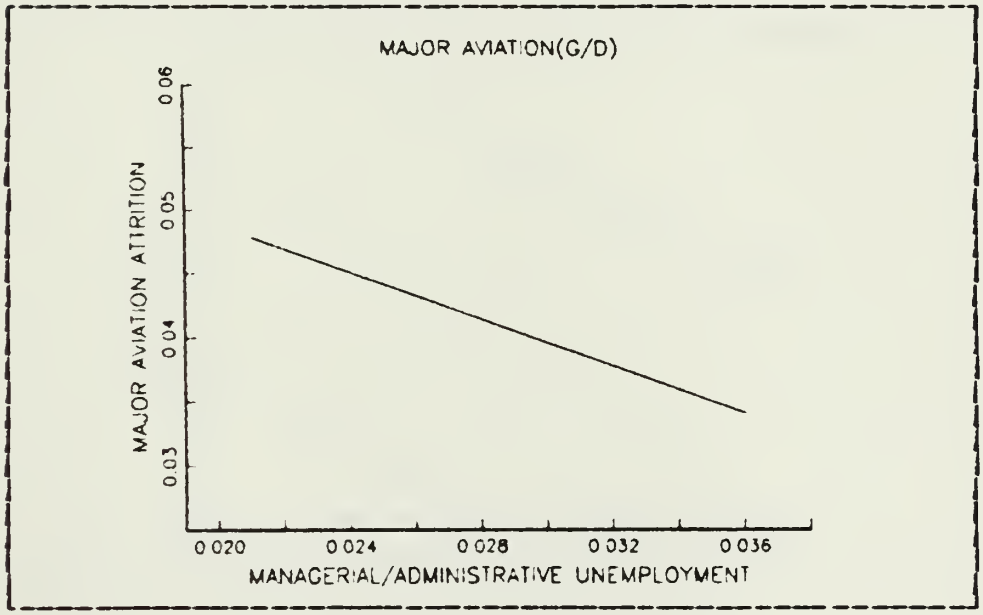


Figure C.7 Major Air/Using General Differencing.

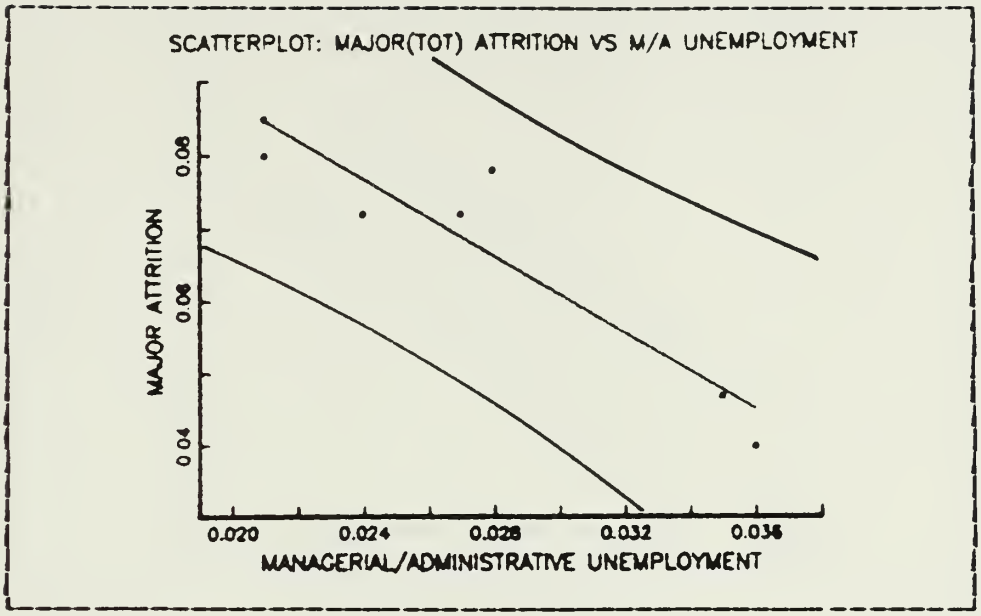


Figure C.8 Major Total.

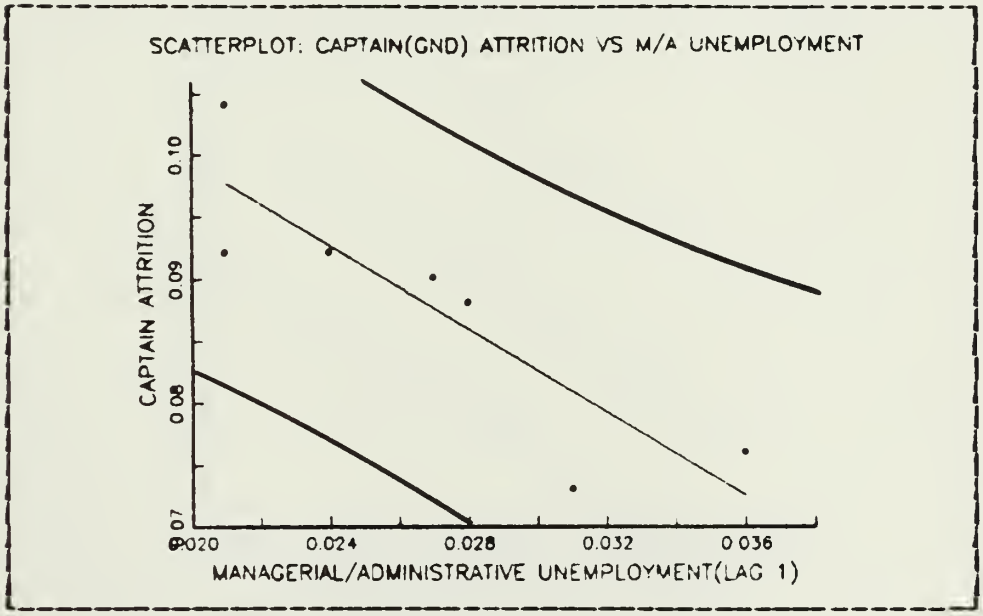


Figure C.9 Captain Ground.

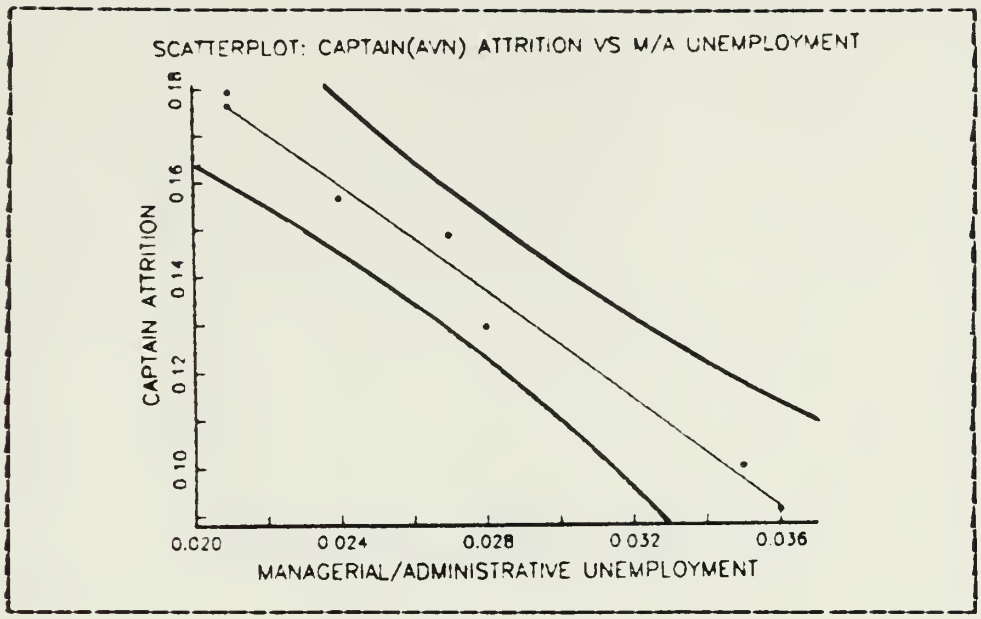


Figure C.10 Captain Aviation.

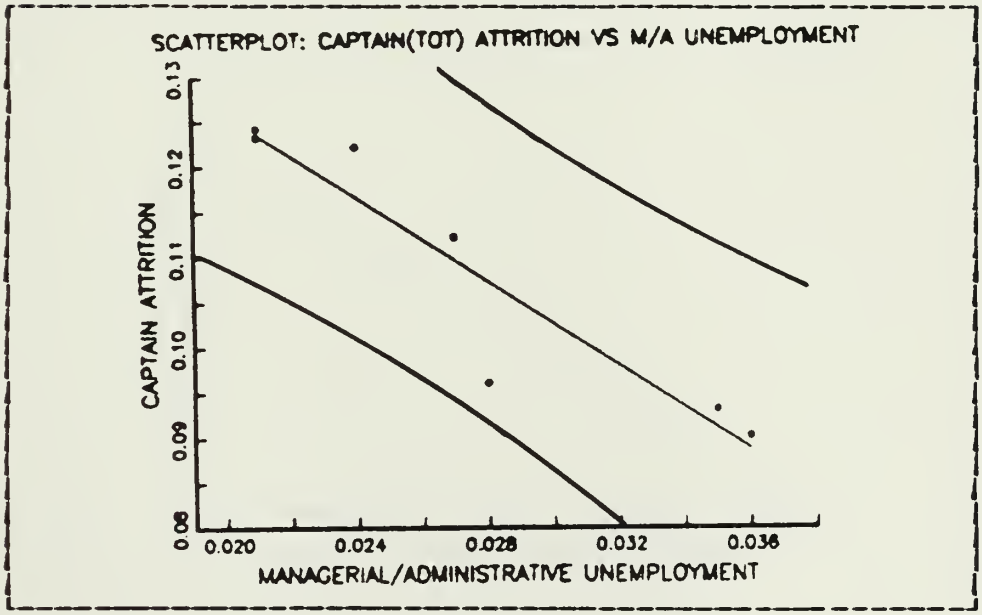


Figure C.11 Captain Total.

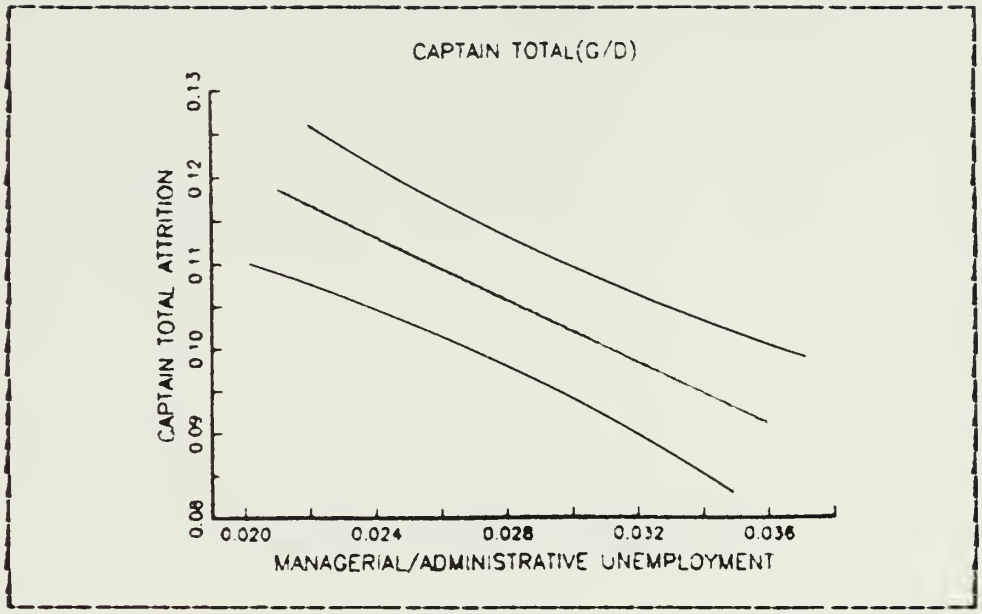


Figure C.12 Captain Total/Using General Differencing.

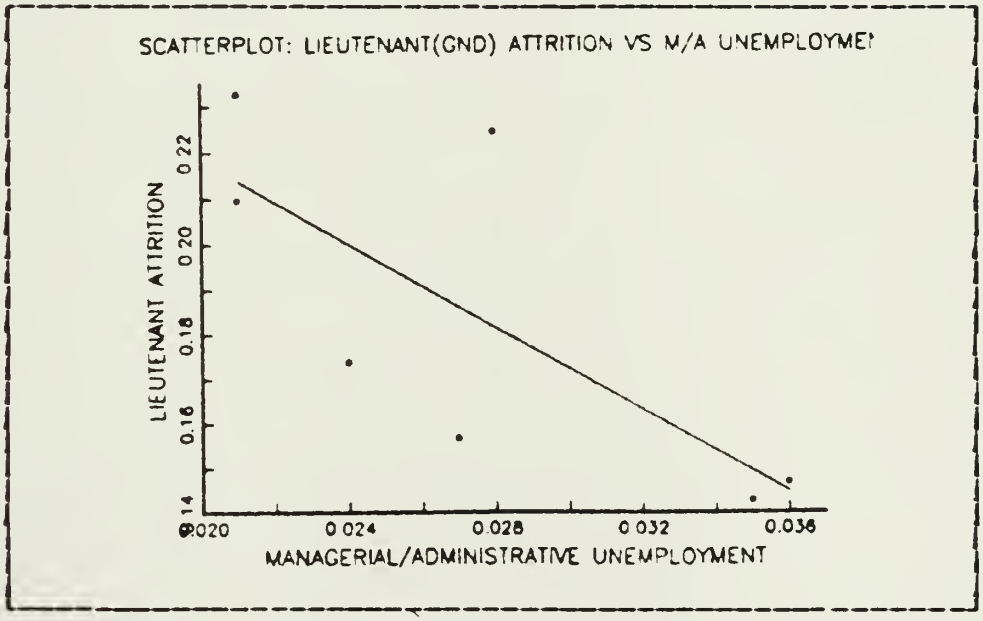


Figure C.13 Lieutenant Ground.

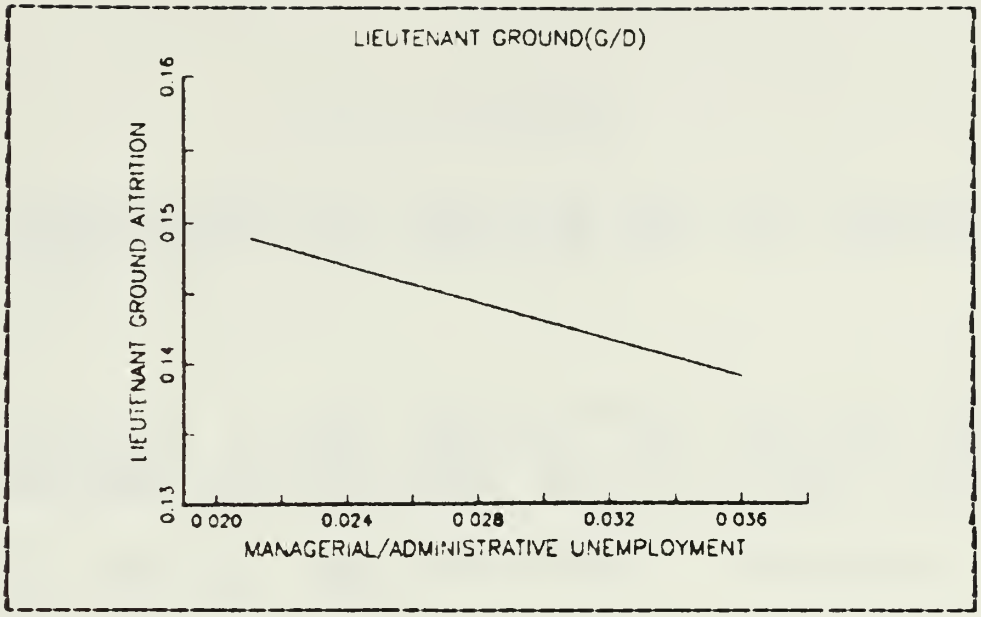


Figure C.14 Lieutenant Ground/Using General Differencing.

IIST OF REFERENCES

1. Center for Naval Analyses 83-3116/22, CNA Memorandum: "Marine Officer Loss Rates for 1984," by Kathleen P. Utgoff, July, 1983.

2. Ibid., pp. 2-3.

3. Green, G. P. and others, "Revisions in the Current Population Survey Beginning in January 1983," Employment and Earnings, v. 30, pp.7-11, U.S. Government Printing Office, Wash., D.C., Nov., 1983.

4. SAS Institute, SAS User's Guide: Statistics, SAS Institute Inc., 1982.

5. Pindyck, Robert S., Econometric Models & Economic Forecasts, pp. 157-158. Irwin, 1981.

6. Navy Personnel Research and Development Center, TR 83-30, "Methods for Forecasting Officer Loss Rates," Siegel, Barry, p. 7, August 1983.

7. Center for Naval Analysis 83-3116/22, CNA Memorandum: "Marine Officer Loss Rates for 1984," by Kathleen P. Utgoff, July, 1983.

8. Ibid., p. 5.

9. Pindyck, Robert S., Econometric Models & Economic Forecasts, pp. 47-48, Irwin, 1981.

10. Ibid., pp. 65-67.

11. The Rand Corporation Memorandum RM-6103-SA, "An Introduction To Cost Estimating," by C. A. Batchelder, pp. 55-56, Rand Corporation, 1969.

12. Ibid., pp. 47-48.

13. Ibid., pp. 47-48.

14. Ibid., pp. 43-44.

15. Pindyck, Robert S., Econometric Models & Economic Forecasts, pp. 152-164., Irwin, 1981.

16. Ibid., pp. 158-160.

17. Ibid., pp. 152-154.

18. Ibid., pp. 157-158.

19. The Rand Corporation Memorandum RM-6103-SA, "An Introduction To Cost Estimating," by C. A. Batchelder, pp. 50-54, Rand Corporation, 1969.

BIBLIOGRAPHY

Box, George E. P., Statistics For Experimenters, Wiley, 1978.

Duncan, Acheson J., Quality Control and Industrial Statistics, F Irwin, 1974.

Green, G. P. and others, "Revisions in the Current Population Survey Beginning in January 1983," Employment and Earnings, v. 30, U.S. Government Printing Office, Feb, 1983.

Michaels, Charles E., and Spector, Paul E., "Causes of Employee Turnover: A Test of the Mobley, Griffeth, Hand, and Meglino Model," Journal of Applied Psychology, vol. 67, No. 1, 53-59, 1983.

Mobley, W. H. and Meglino, B. M., "Review and Conceptual Analysis of the Employee Turnover Process," Psychological Bulletin, vol. 86, No. 3, 493-522, 1979.

Mobley, W. H., "Intermediate Linkages in the Relationship Between Job Satisfaction and Employee Turnover," Journal of Applied Psychology, vol. 62, no. 2, 237-240, 1977.

Youngblood, Stuart A., and Mobley, William H., "A Longitudinal Analysis of the Turnover Process," Journal of Applied Psychology, Vol. 68, No. 3, 507-516, 1983.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943	2
3.	Professor Dan Boger, Code 54Bk Naval Postgraduate School Monterey, California 93943	1
4.	Professor Don Barr, Code 55Bn Naval Postgraduate School Monterey, California 93943	1
5.	Professor Robert R. Read, Code 55Re Naval Postgraduate School Monterey, California 93943	1
6.	LTC M. Hagee HCMC, Code MPP Washington, D.C. 22314	1
7.	MAJ W. J. Esmann 4509 Banf Street Annandale, VA 22314	1

route to:GENReed, Dorothy J.

ID:32768001014756
E686
Marine officer attrit
\Esmann, William John.
due:11/7/1996,23:59

ID:32768000999270
0246
An analysis of West P
\Oehrlein, Richard Vin
due:11/7/1996,23:59

ID:32768000768816
G8333
The effects of econom
\Grimes, Gary R.
due:11/7/1996,23:59

ID:32768003196718

428347

Thesis
E686 Esmann
c.1 Marine officer attri-
tion model.

AMERICAN BINDERIES

thesE686
Marine officer attrition model.
3 2768 001 01475 6
DUDLEY KNOX LIBRARY